Dynamics of the innate immune response to infection: a mathematical modelling approach

Sophie Vanessa Hiu-Lam Ip

This thesis is submitted for the degree of Doctor of Philosophy

Department of Applied Mathematics and Theoretical Physics

Jesus College

University of Cambridge

June 2021

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface (below) and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

The ASC cerulean datasets (FLICA and MitoTracker) presented in Chapter 2 were created by Dr Alessandro Rizzo. The methodology presented in Chapter 2 for extracting traces from these datasets was based on Matlab code written by Dr Alessandro Rizzo and Dr Eugenia Cammarota.

Dynamics of the innate immune response to infection: a mathematical modelling approach

Sophie Vanessa Hiu-Lam Ip

The innate immune system plays a vital role in the control of infection, but many aspects of its behaviour are not fully understood. In this dissertation, I use mathematical modelling approaches to elucidate two linked processes which are key to the innate immune response to *Salmonella* infection.

First, I consider the inflammasome, a large protein complex which forms in cells following infection. The inflammasome is responsible for coordinating cell death and the release of cytokines, which promote further activation of the immune system. Despite its importance, there are many conflicting accounts of this process in the existing literature, and there is no accepted single conceptual model of inflammasome formation.

In the first chapters of this dissertation, I present a suite of deterministic ordinary differential equation models and discrete stochastic models which highlight different elements of the inflammasome formation process. By comparing the results from these models with an existing dataset describing the innate immune response to infection in macrophages, I construct a cohesive conceptual model of inflammasome formation. In particular, I propose a novel 'branching' mechanism of protein recruitment to the inflammasome complex. I also show that variation in inflammasome formation times can arise from differences in the initial abundance of NLR oligomers present in the cell prior to inflammasome formation.

The remainder of this dissertation focuses on the formation of clusters of infected cells (or

'lesions') in tissues following *Salmonella* infection. Inflammasome formation on the cellular level is a crucial part of this larger-scale process; however, the precise downstream effects of inflammasome-coordinated signalling and cell death in this context remain unclear.

In the later chapters of this thesis, I outline a spatial partial differential equation model of lesion formation in liver tissue following *Salmonella* infection, with a particular focus on motility of infected cells and multiple forms of cell death. Through use of Turing instability analysis and simulated solutions of the lesion model, I show that the spatial structure of lesions can arise simply through the interactions of infected and uninfected phagocytes, bacteria and a chemokine. I also demonstrate the importance of decreased motility of infected cells, and a balance of different forms of cell death with influx of uninfected immune cells, for the successful control of infection.

Acknowledgements

My thanks are due, first and foremost, to my supervisors, Professor Julia Gog and Professor Clare Bryant. I would not have started this PhD without their encouragement, nor would I have finished it without their wisdom, support and patience. It has been an honour to work with both of them, and I am deeply grateful for all they have done for me.

I am also extremely grateful to the other members past and present of the Disease Dynamics Group at DAMTP, especially Veni Karamitsou, Stephen Kissler, Petra Klepac, Jordan Skittrall, and Maria Tang, for keeping me company on the journey of this PhD, and for their invaluable mathematical advice.

Likewise, my thanks go to the members of the Bryant Lab, at the Vet Schools and Medicine Department, for giving me an academic home-away-from-home. Thanks are especially due to Lee Hopkins, Charlotte Macleod and Zsofia Digby, for helping a mathematician find her feet in the world of experimental biology.

I would also like to acknowledge the support of the BBSRC DTP, which provided funding for this PhD, as well as numerous opportunities for training and personal development, and administrative assistance.

Jesus College has also been an essential lifeline throughout the PhD process. Thanks to everyone there who has played a role in enriching my time here in Cambridge, especially the folks at Jesus Chapel.

Finally, to all the friends and family who have made this PhD possible: thank you. In particular, thanks to James Wilkinson for being endlessly supportive, funny, smart and interesting, and to Em Black and Nate Dunmore, who are the most caring (and entertaining) housemates I could wish for. Thanks to my parents and Caroline, who have always encouraged me and been there for me, through thick and thin. And finally, thank you to Robbie, for his generosity, for his thoughtfulness, for his gentleness, for everything.

Contents

List of Figures

List of Tables

List of abbreviations

ABM: Agent-based model

AIC: Akaike information criterion

AIM: Absent in melanoma

BD: Becker-Döring

CAPS: Cryopyrin-associated periodic syndrome

CARD: Caspase recruitment domain

COP: CARD-only protein

CTMC: Continuous time Markov chain

ECM: Extracellular matrix

EM: Electron microscopy

GLM: Generalised linear model

GSDMD: Gasdermin-D

IC: Initial condition

IFN-γ: Interferon γ

IKK: IκB kinase

IL: Interleukin

LPS: Lipopolysaccharide

LRRK: Leucine-rich repeat-containing kinase 2

MARK4: Microtubule-affinity regulating kinase 4

MAS: Macrophage activation syndrome

ML: Marcus-Lushnikov

NAIP: Neuronal apoptosis inhibitory protein

NEK7: NIMA-related kinase 7

NLR: NOD-like receptor

NOD: Nucleotide-binding oligomerisation domain

NP: Nucleated polymerisation

NF-κB: Nuclear factor-κ B

ODE: Ordinary differential equation

PAMP: Pathogen-associated molecular pattern

PDE: Partial differential equation

PKCδ : Protein Kinase C δ

POP: Pyrin-only protein

PRR: Pattern recognition receptor

PTM: Post-translational modification

PYD: Pyrin domain

RNS: Reactive nitrogen species

ROI: Region of interest

ROS: Reactive oxygen species

T3SS: Type III secretion system

TLR: Toll-like receptor

TNF-α: Tumour necrosis factor α

YFP: Yellow fluorescent protein

Chapter 1

Introduction

1.1 The role of innate immunity in response to infection

The innate immune system is the human body's first line of defence against infection. Its basic roles in this context are the identification of pathogens, the clearance of infection, and the promotion of further immune system activation - the latter may involve the recruitment of other immune cells, the promotion of inflammation by the release of chemical signals known as cytokines, and the activation of the adaptive immune system. The regulatory systems which underlie the response to infection are complex, and rely on the coordination of processes from the subcellular level through tissues, organs and organ systems to the organism as a whole. Understanding how these processes function, and in particular, the link between these scales, is key to understanding how the clearance of infection is managed.

This dissertation will focus on the response to *Salmonella enterica* infection. *Salmonella* is responsible for life-threatening diseases in humans and animals, including enteric fevers (with typhoidal serotypes of *Salmonella* responsible for typhoid fever in humans) and gastroenteritis. Enteric fevers are a particular cause for concern in human health, with around 14 million people affected per year, mainly in lower-income countries [Stanaway et al., 2019]. Thus, understanding how the innate immune system contributes to the clearance of *Salmonella* infection, and how this process can malfunction, is extremely valuable; moreover, understanding how *Salmonella* is cleared will also give insight into innate immune system more generally.

In particular, I will be examining two steps in the innate immune response to *Salmonella* infection which occur on different timescales: inflammasome formation, a subcellular pro-

cess which occurs on a timescale of minutes following the infection of an individual cell, and the formation of lesions in the liver, which occurs on a timescale of days.

Broadly, there are two interconnected processes which coordinate response to infection at the cellular level. Firstly, receptors on the cell surface such as Toll-like receptors (TLRs) may detect the extracellular presence of bacteria, leading to the transcription of pro-inflammatory cytokines and the upregulation or post-translational modification (PTM) of proteins involved in pro-inflammatory signalling pathways [Bauernfeind et al., 2009, Brubaker et al., 2015, Swanson et al., 2019]; the latter processes prepare the cell for further responses at a later stage. Secondly, bacteria are phagocytosed, but may survive and reproduce in a vacuole within the cell, shielded from the extracellular effects of the immune system; this is critical for the virulence of *Salmonella* [Fields et al., 1986]. However, innate immune cells have evolved a tightly-mediated signalling cascade to counter this strategy of pathogen survival within the cell, which initiates cell death and the release of pro-inflammatory cytokines (including chemokines which attract other immune cells to the infection site), as well as reducing the motility of the cell [Stehlik et al., 2003b, Man et al., 2014a]. This signalling system is coordinated by the inflammasome, a supramolecular complex which forms following infection and is responsible for processing effector proteins and cytokines. Inflammasomes are also involved in the response to intracellular detection of numerous other threats, either pathogenic or related to cellular damage, and are therefore a key element of innate immune signalling. Inflammasome dysregulation has also been associated with numerous conditions including cancer [Lee et al., 2019, Tartey and Kanneganti, 2019, Velloso et al., 2019, Zhen and Zhang, 2019], neurodegenerative diseases [Boxberger et al., 2019, Han et al., 2019], and autoinflammatory conditions [Canna et al., 2014, Shin et al., 2019, Boxberger et al., 2019]. Despite its importance, there are many open questions surrounding the purpose, structure and formation process of inflammasomes, many of which are highly debated in the existing literature.

However, naturally, within-cell processes are not the end of the story. The signalling pathways at the single-cell level ultimately lead to downstream effects which have ramifications for much more than just the fate of the cell in question; the release of chemokines and reduction of cellular motility are particularly relevant here. It is not fully clear how these downstream processes contribute to the clearance of infection on a larger scale. The effects of *Salmonella* at a tissue level are well-documented; in particular, during systemic infec-

tions, infected cells in the liver and spleen become contained within foci known as lesions. Once again, there are many open questions surrounding lesion formation, and their precise role in limiting the spread of infection is debated. In particular, it is unclear how single-cell processes such as inflammasome formation give rise to lesion formation at the tissue level.

The overall aim of the research presented in this dissertation is to shed light on these two key elements in the innate immune response, and clarify the processes which are taking place. There are two fundamental questions I wished to answer: namely, how do inflammasomes form? And why do lesions form? The results presented in the following chapters may go some way towards clarifying these issues. Both questions are also linked to an implicit third question: how does one link an understanding of small scale processes (for example, structural studies of interactions between proteins involved in the inflammasome, or the actions of individual cells within a lesion) to an understanding of a larger-scale process which arises from a large number of smaller scale processes (the formation of the inflammasome or lesion as a whole)? In order to address these questions, I have used a mathematical modelling approach. This is ideally suited to the third unifying question, since quantitative analysis can give an insight into how many individual constituent parts in a system give rise to an overall outcome. Mathematical modelling has not been particularly widely used in the study of innate immunity, with the majority of work in this field being experimental in nature, although in recent years the use of high-throughput technologies has necessitated the introduction of statistical and computational approaches. However, the more abstract approach of mathematical modelling, focusing on conceptual model creation and translation into mathematical systems, has been much rarer. There have been some previous studies modelling the formation of lesions (these will be further explained in Section 1.7) but these have not been particularly numerous, and do not focus on the scale-crossing aspects of this process; models of inflammasome formation are even scarcer. As well as providing useful and interesting results in their own right, it is my hope that the models presented in this dissertation will motivate a repeating cycle of experimental work and mathematical modelling, as follows: experimental work inspires a model by providing data and highlighting areas of a biological system which are not well-understood; a model is then developed which may suggest explanations for the behaviours observed by experimental work, as well as posing questions to be addressed through further experimental work. Experiments should then be conducted to verify the results of the modelling work and address the questions posed therein. However,

this experimental work will inevitably bring to light even more areas which require further investigation, and thus the cycle repeats, with models and experiments becoming increasingly refined. The work presented in this dissertation represents a first iteration of this cycle.

1.2 Structure of this dissertation

This dissertation broadly encompasses two separate but thematically-linked sections, with the aims of modelling the inflammasome and lesion formation respectively.

In this introductory chapter, I will describe the relevant existing literature for modelling both inflammasomes and lesions, from both biological and mathematical perspectives, as well as outlining the research aims.

The first chapters of this dissertation are dedicated to inflammasome modelling; in Chapter 2 I present an existing experimental dataset which motivates this research, and in Chapter 3 I present two novel mathematical models of inflammasome formation. Chapters 4 and 5 are dedicated to analysis of these models, and Chapter 6 links this analysis back to the experimental data presented in Chapter 2. Finally, in Chapter 7, I discuss the results of the inflammasome modelling project.

I then turn my attention to the lesions modelling project. In Chapter 8 I present a novel model of lesion formation, which I analyse using analytical and computational techniques in Chapters 9 and 10. Finally, in Chapter 11 I discuss the results of the lesions modelling project.

The dissertation concludes with a discussion of the broader implications of this research, as well as suggesting avenues for further work, both mathematical and experimental.

1.3 The biology of inflammasome formation

Despite its importance, the process of inflammasome formation is not well understood. There is a wealth of existing literature describing inflammasome formation on a structural level; the domain architecture of inflammasome constituent proteins have been mapped, and the modes of pairwise interactions between individual proteins has been documented. The effects of inflammasome formation are also well-understood on the cell-wide scale; the levels of inflammasome formation can easily be measured either by quantifying levels of out-

put of cytokines or cell death (the downstream products of inflammasome formation) or by observing the presence of inflammasome structures using microscopy techniques.

However, it is unclear how the system behaves between these two scales; in particular, how combinations of the well-documented protein-protein interactions give rise to complex structures containing thousands of individual proteins. Few experiments have been successful in exploring the region between these scales, although this is currently an area of considerable interest. Some confocal imaging experiments have established mid-scale observations about inflammasome structure, and there has also been some success with cryo-electron microscopy (EM) techniques, but these approaches are all in their early stages. Conceptual models have been proposed for the internal architecture of the inflammasome structure, but these are often vague, are based on numerous assumptions, and have not been properly motivated by experimental observation.

In the first section of this chapter, I will give an overview of the relevant existing literature surrounding inflammasome formation on these distinct 'micro' (structural) and, comparatively, 'macro' (whole cell-level) scales, as well as discussing previous attempts to unify these scales to create a cohesive conceptual model of inflammasome formation.

1.3.1 The role of the inflammasome

The formation of the inflammasome is triggered by the activation of cytosolic pattern recognition receptors (PRRs), which occurs following detection of pathogen- or damage-associated molecular patterns (PAMPs or DAMPs). PRRs then oligomerise and recruit a scaffold protein which forms the bulk of the complex, which may also act as an adaptor [Fernandes-Alnemri et al., 2007]. Finally, the scaffold protein recruits and processes a final effector protein. These three classes of protein (PRR, adaptor and effector) together form the inflammasome.

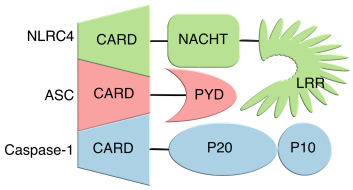
Different PRRs are activated depending on the PAMP/DAMP; the PRRs involved in inflammasome formation are members of the NOD (nucleotide-binding oligomerisation domain)-like receptor (NLR) family, AIM2 (absent in melanoma 2) and pyrin, with different PRRs responding to different PAMPs/DAMPs. Here we focus on the two most-studied NLRs, NLRC4 and NLRP3, since these are the best-characterised inflammasomes; however, many of the approaches explored in this thesis could be applicable to inflammasomes based on other PRRs.

NLRC4 is extremely specialised in its function; it is responsible only for the intracellular

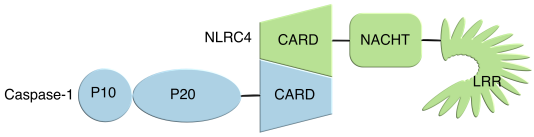
detection of Gram negative bacteria [Zhao et al., 2011]. However, its structure and mode of oligomerisation have been extremely well studied. On the other hand, NLRP3 responds to a wide variety of cellular threats, including bacterial, viral and fungal infections, as well as various damage-associated signals and foreign irritants [Kanneganti et al., 2006, Warren et al., 2008, Thomas et al., 2009, Allen et al., 2009, Gross et al., 2009]. A comprehensive table of NLRP3 stimuli can be found in a recent review [Swanson et al., 2019]. The NLRP3 inflammasome has also been the focus of much experimental work since it can be activated without the use of pathogens; in particular, a common method is to 'prime' the cell with lipopolysaccharide (LPS) to upregulate NLRP3 and apply post-translational modifications [Bauernfeind et al., 2009, Swanson et al., 2019]; this is followed by treatment with nigericin to trigger inflammasome formation. In contrast, a common method for triggering NLRC4 inflammasome formation is to infect cells directly with *Salmonella*, often using fluorescently labelled bacteria so that the infection state of the cell can easily be determined. However, the mode of activation and oligomerization of NLRP3 are considerably more complicated than those of NLRC4, and in particular, the structure of the NLRP3 oligomer has not been studied in detail.

For inflammasomes including either of these proteins, the scaffold protein is apoptosis-associated speck-like protein containing a CARD (ASC), which provides a platform for caspase-1 processing via proximity-induced proteolytic activation [Stehlik et al., 2003b, Sanders et al., 2015, Schmidt et al., 2016]. Caspase-1 is an effector protein responsible for the processing of pro-inflammatory cytokines interleukin 18 (IL-18) and IL-1β, as well as gasdermin-D (GSDMD) [Srinivasula et al., 2002, Dinarello, 2009]; following cleavage by caspase-1, the N-terminus of GSDMD creates pores in the cell membrane, allowing cytokines to exit the cell, and ultimately leading to pyroptosis, a lytic and pro-inflammatory form of cell death [Shi et al., 2015a, Liu et al., 2016, Sborgi et al., 2016, Evavold et al., 2018, He et al., 2015]. It has also been suggested that ASC specks are released from cells during pyroptosis and continue to process cytokines, or may be taken up by other cells to initiate further ASC speck formation; however, these theories are controversial [Franklin et al., 2014, Baroja-Mazo et al., 2014].

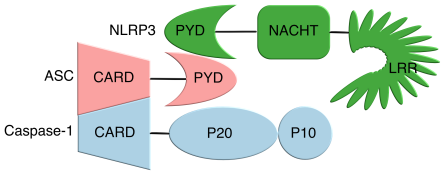
So-called 'non-canonical' inflammasomes, may also be activated by caspases other than caspase-1, including caspase-4 and caspase-5 in humans [Baker et al., 2015], and caspase-11 in mice [Kayagaki et al., 2011]; this leads to pyroptosis and cytokine release via secondary activation of the NLRP3 inflammasome.



Interactions in the NLRC4 inflammasome; NLRC4 recruits ASC via CARD-CARD interactions, and ASC recruits caspase-1 via CARD-CARD interactions



Alternative characterisation of the NLRC4 inflammasome; NLRC4 recruits caspase-1 directly via CARD-CARD interactions



Interactions in the NLRP3 inflammasome; NLRP3 recruits ASC via PYD-PYD interactions, and ASC recruits caspase-1 via CARD-CARD interactions

Figure 1.1: Schematic of protein-protein interactions in the NLRC4 and NLRP3 inflammasomes; reactions have been simplified to be represented in two dimensional view.

1.3.2 Structural properties

Protein structures

NLRC4 and NLRP3 are both tripartite proteins. Both contain an N-terminal domain which recruits ASC: a CARD (caspase activation and recruitment) domain in the case of NLRC4, and a PYD (pyrin) domain in the case of NLRP3 [Oroz et al., 2016]. They also contain a central nucleotide binding NACHT domain, and a C-terminal LRR domain [Ting et al., 2008]. The three domains are joined by flexible linkers. ASC consists of an N-terminal PYD domain and a C-terminal CARD domain, also joined by a flexible linker [de Alba, 2009]. Finally, caspase-1 consists of an n-terminal CARD domain joined by a flexible linker to P20 and P10 subunits; these are cleaved following activation [Martinon et al., 2002, Manji et al., 2002, Boucher et al., 2018] . These structures are illustrated in Table 1.1. Proteins from these families associate via homotypic PYD-PYD and CARD-CARD interactions; NLRP3 and NLRC4 recruit ASC via PYD-PYD and CARD-CARD interactions respectively, and ASC then recruits caspase-1 via CARD-CARD interactions (Figure 1.1) [Masumoto et al., 2001, Srinivasula et al., 2002, Stehlik et al., 2003b, Sahillioglu et al., 2014, Sborgi et al., 2015].

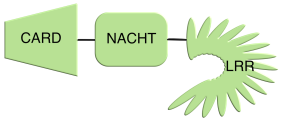
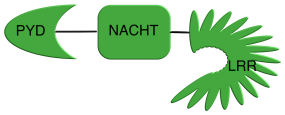
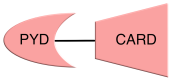
NLRC4	
NLRP3	
ASC	
Caspase-1	

Table 1.1: Schematic structure of proteins involved in inflammasome formation

NLRs in the inflammasome

NLRs are generally present throughout innate immune cells in the absence of PAMPs/DAMPs, albeit in an autoinhibited form in which they cannot interact with other inflammasome constituents. In the autoinhibited form of NLRC4 derived by Hu *et al.*, the LRR hinges down to occlude the other domains [Hu et al., 2013]. The precise mechanism of NLRP3 inhibition remains unclear; while some studies suggest that the LRR is also responsible for autoinhibition in NLRP3, this remains controversial [Hu et al., 2013, Hafner-Bratkovič et al., 2018, Platnich and Muruve, 2019, Swanson et al., 2019]. For inflammasome formation to take place, NLRs must be converted into an active form and oligomerize to 'seed' the inflammasome. This activation and oligomerization process is controlled by a complex system of interacting regulatory processes, including gene expression, PTMs, and ligand binding.

Prior to cellular infection, NLRP3 expression is upregulated by pro-inflammatory stimuli which themselves are induced by other pathways; other PRRs such as TLRs and nucleotide-binding oligomerization domain-containing protein 2 (NOD2), as well as cytokines such as tumour necrosis factor α (TNF-α) and IL-1β induce increased expression of NLRs via activation of transcription factor nuclear factor κ B (NF-κB) [Franchi et al., 2009, Bauernfeind et al., 2009]. NF-κB also induces expression of IL-1β which is not constitutively expressed in macrophages [Bauernfeind et al., 2009]. These priming signals do not affect levels of ASC, pro-IL-18 or pro-caspase-1 [Bauernfeind et al., 2009].

However, interestingly, a basal level of NLRC4 is sufficient for inflammasome formation, whereas basal NLRP3 levels are not. It has been shown that various microRNAs are responsible for ensuring low levels of NLRP3 expression [Zamani et al., 2020, Boxberger et al., 2019], although they may also be implicated in NLRP3 activation [Ojcius et al., 2019]. An initial stage of cell 'priming' is required for sufficient levels of NLRP3 expression for inflammasome formation to be possible [Bauernfeind et al., 2009], except in the case of a few 'alternative inflammasome' pathways [He et al., 2013, Gaidt et al., 2016, de Carvalho et al., 2019].

Following upregulation, various PTMs are necessary for inflammasome activation. In NLRC4, some studies suggest that phosphorylation of residue Ser533 occurs following exposure to Salmonella and is necessary for endogenous levels of IL-1β production and cell death following infection. It has been suggested that protein kinase C δ (PKCδ) or leucine-rich repeat-containing kinase 2 (LRRK2) may be responsible, as both have been shown to

associate with phosphorylated NLRC4, and depletion of either protein results in reduced NLRC4 inflammasome activation [Qu et al., 2012, Matusiak et al., 2015, Liu et al., 2017]. However, the importance of Ser533 is debated [Tenthorey et al., 2020]; furthermore, it is not clear what precise effect this phosphorylation may have on NLRC4 activation. It is possible that phosphorylation of Ser533 may have an important role in stabilising NLRC4 in its active form, but is insufficient for full NLRC4 activation alone.

The picture of regulation of NLRP3 via PTMs is somewhat more complicated; to date at least 13 phosphorylation, ubiquitination and sumoylation sites have been identified in NLRP3, some of which activate, and some of which inhibit or have no effect on inflammasome activity [Kelley et al., 2019, Swanson et al., 2019, Yang et al., 2019]. It has been suggested that priming via PTMs prior to activation is a necessary prerequisite for full activation, although it is currently unclear how these PTMs are regulated.

Subcellular location of NLRP3 during inflammasome activation has also been shown to play a significant role. Different studies appear to show conflicting results, although there is a general consensus that during inflammasome formation, NLRP3 colocalises to the Golgi and endoplasmic reticulum near the nucleus. Mitochondrial clustering and association with NLRP3 has also been observed [Zhou et al., 2011, Misawa et al., 2013, Li et al., 2017]. It has been demonstrated that NLRP3 associates with microtubule-affinity regulating kinase 4 (MARK4), which drives its movement along microtubules to the microtubule-organizing centre [Li et al., 2017]; disruption of MARK4 decreases, but does not completely abrogate inflammasome formation. Neither movement along microtubules nor mitochondrial clustering appear in association with NLRC4 inflammasome formation, although it has been shown that actin polymerisation is necessary for inflammasome formation [Man et al., 2014a], suggesting a role for the actin cytoskeleton in NLRC4 inflammasome formation which may be similar to that of the microtubule network in NLRP3 inflamamsome formation.

A final method by which NLR oligomerization is regulated is interaction with pyrin-only proteins and CARD-only proteins (POPs and COPs respectively). These proteins, which are present only in primates, bind to NLRs via PYD-PYD and CARD-CARD interactions, so that these domains are unavailable for interaction with other inflammasome constituents. These proteins therefore act as a negative regulatory mechanism for inflammasome formation [Stehlik et al., 2003a, Dorfleutner et al., 2007a, Dorfleutner et al., 2007b, Atianand and Harton, 2011]. NF-κB and IL-1β both upregulate the expression of POP1 and POP2, sug-

gesting a negative feedback mechanism [Atianand and Harton, 2011, Bedoya et al., 2007]. However, it remains unclear how any of these regulatory processes are themselves controlled, and to what extent they are important.

These regulatory processes in isolation are insufficient to prompt full NLR activation. In the case of NLRC4, binding of a PAMP is necessary; this is an indirect interaction via an intermediary neuronal apoptosis inhibitory protein (NAIP). Four NAIPs have been identified in mice, although only one is present in humans. It has been demonstrated that during *Salmonella enterica* serovar Typhimurium infections in mouse cells, different NAIPs recognise different *Salmonella*-associated PAMPs. NAIP1 recognises PrgI and NAIP2 recognises PrgJ, both constitutents of the type III secretion system (T3SS); NAIP5 and NAIP6 both recognise flagellin [Lightfield et al., 2011, Zhao et al., 2011, Yang et al., 2013, Tenthorey et al., 2014]. In human cells, the single human NAIP recognises multiple PAMPs, including elements of the T3SS in multiple species of bacteria, as well as flagellin from Salmonella. In all cases, these interactions take place via ligand binding, changing the conformation of the NAIP involved so that it may in turn form a complex with, and activate, NLRC4 [Hu et al., 2015, Zhang et al., 2015, Tenthorey et al., 2017, Yang et al., 2018].

The activation of NLRP3 is somewhat more complex. NLRP3 does not interact directly with the many PAMPs and DAMPs it recognises, but instead appears to be affected by the downstream effects of the PAMP or DAMP in question on the intracellular environment. No single activating effect has been isolated, but links have been established with processes including potassium and chloride ion efflux from the cell [Verhoef et al., 2005, Muñoz-Planillo et al., 2013], calcium ion flux [Murakami et al., 2012], the production of reactive oxygen species (ROS) [Cruz et al., 2007, Dostert et al., 2008], lysosomal disruption [Hornung et al., 2008], and mitochondrial dysfunction [Zhou et al., 2011, Shimada et al., 2012]. In particular, potassium ion efflux is often cited as a putative underlying common feature to these conditions, although it has been shown that NLRP3 inflammasome formation can occur independently of potassium ion efflux [Meng et al., 2009, Groß et al., 2016, Sanman et al., 2016]. It is possible that NLRP3 inflammasome formation is in fact activated combinatorially by these different cellular conditions, with high overall levels of one or more such factors leading to activation.

Notably, spontaneous inflammasome activation appears in a number of autoinflammatory conditions including macrophage activation syndrome (MAS), and cryopyrin-associated

periodic syndromes (CAPS). Both of these conditions are caused by mutations in the genes encoding NLRC4 and NLRP3 respectively, which cause spontaneous oligomerization and thus inflammasome activation in the absence of DAMPs/PAMPs [Meng et al., 2009, Canna et al., 2014, Moghaddas et al., 2018].

Following activation by interaction with a PAMP/DAMP, NLRs may then propagate their active forms to other NLRs still in their autoinhibited form via complex formation. Activation of a single NLR therefore precipitates a chain reaction of NLR activation, and the formation of an oligomer consisting of active NLR monomers. In the case of NLRC4, the formation of this oligomer has been studied through the use of cryo-EM. The structure of a NAIP5-NLRC4 complex has been determined by Zhang *et al.* [Halff et al., 2012, Zhang et al., 2015], and the structure of a NAIP2-NLRC4 complex has been determined by Hu, Zhang and others [Hu et al., 2015, Zhang et al., 2015]; notably both the Hu and Zhang studies use NLRC4 with the CARD domain removed. Active NLRC4 monomers form complexes via interactions between NACHT domains. However, the precise form of the NLRC4 oligomer remains controversial. Many early studies suggested that active NLRC4 forms a closed ring containing 10-12 NLRC4 monomer subunits, and a single NAIP (Figure 1.2); however, other studies suggest that the oligomer may form an extended filamentous structure, with the CARD domains forming a central core [Diebolder et al., 2015, Li et al., 2018] (Halff *et al.* propose that NLRC4 in isolation forms helices, while it forms closed rings when associated with NAIP5 [Halff et al., 2012]). It is possible that earlier studies which show a closed ring oligomer structure may in fact have observed a short segment of a helical oligomer, which has been 'flattened' into a closed ring through errors in averaging. However, this debate has not been satisfactorily resolved, especially as these studies have for the most part been conducted with NLRs and NAIP in isolation, and may not accurately reflect conditions in the cell.

Far less is known about the structure of NLRP3 oligomers than NLRC4 oligomers; it has been suggested that NLRP3 forms rings through ATP-dependent interaction of NACHT domains, although this is a conjecture based on analogy with NLRC4 interactions [Lechtenberg et al., 2014, Sharif et al., 2019]. However, the binding of NIMA-related kinase 7 (NEK7) has been shown to be necessary for NLRP3 inflammasome formation [He et al., 2016, Schmid-Burgk et al., 2016, Shi et al., 2016a]; it appears to bind directly to NLRP3 and facilitate NLRP3 oligomerization, stabilising interactions between adjacent monomers [Sharif et al.,

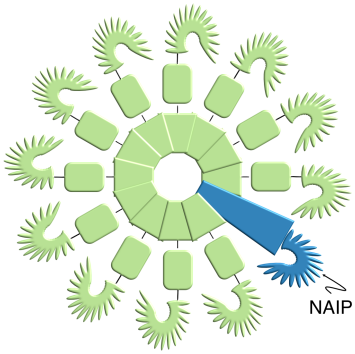


Figure 1.2: A common conceptual model of inflammasome formation (see [Halff et al., 2012, Hu et al., 2015, Zhang et al., 2015]) suggests that the NLRC4 oligomer consists of a closed ring of 10-12 NLRC4 monomers and a single NAIP which interact via NACHT domains.

2019]. Alternative splicing has also been proposed as regulatory mechanisms for NLRP3, with one splice variant lacking the interaction surface necessary for NEK7 binding, and therefore inhibiting oligomerization [Hoss et al., 2019].

Recent studies have demonstrated that other inflammasome PRRs including NLRP1 and AIM2 oligomerise into helical structures [Gong et al., 2021, Hollingsworth et al., 2021, Matyszewski et al., 2021] using cryo-EM techniques; however, it is worth noting that these studies focus on the behaviour of individual domains (PYD in the case of AIM2, and CARD in the case of NLRP1). It is therefore worth bearing in mind that the structures formed by the full-length proteins may be slightly different.

ASC in the inflammasome

Similarly to NLRC4 and NLRP3, ASC is present endogenously throughout innate immune cells. Following PAMP/DAMP detection, it is recruited into the inflammasome by PYD-PYD or CARD-CARD interactions with NLRP3 and NLRC4 respectively, and further aggregates around this NLR seed to form a well-defined focus or 'speck'. The speck is around 0.7 μm in diameter and can be easily visualised on an individual cell level by fluorescently labelled ASC [Cheng et al., 2010, Man et al., 2014b, Kuri et al., 2017].

The most basic assumption for the role of ASC is that it acts as an adaptor protein, forming a bridge between NLRs and caspase-1. In the case of the NLRP3 inflammasome this is certainly the case. In the NLRC4 inflammasome, however, caspase-1 may interact directly with NLRC4 via CARD-CARD interactions, although ASC specks are still observed in infected cells, and the absence of ASC has been shown to lead to significantly decreased

levels of NLRC4 inflammasome activation [Stehlik et al., 2003b, Proell et al., 2013, Li et al., 2018]. It is also unclear on the basis of this reasoning why the formation of such a large structure would be necessary if ASC is merely acting as an intermediary between NLRs and pro-caspase1.

Once ASC has been recruited by NLRs, a common conceptual model of inflammasome formation suggests that ASC forms filaments via PYD-PYD interactions, since ASC PYD domains form filamentous structures in isolation, via homotypic interactions [Vajjhala et al., 2014, Lu et al., 2014, Sborgi et al., 2015]. However, this might suggest a stellate form for the completed inflammasome, with multiple ASC filaments projecting from an NLR core. This does not reflect the inflammasome phenotype in endogenous cells, which generally appear as 'smooth', compact structures. Some mutant phenotypes do display a somewhat stellate or 'Medusa's head' appearance, both in ASC overexpression systems, and systems in which ASC CARD-CARD interactions have been disrupted [Masumoto et al., 2001, Sahillioglu et al., 2014]. Mutants expressing only ASC PYD domains show formation of filamentous structures, whereas mutants expressing only ASC CARD domains show formation of 'punctate' structures [Sahillioglu et al., 2014, Dick et al., 2016, Kuri et al., 2017].

One proposed explanation for these different phenotypes is that the PYD-PYD ASC filaments are later crosslinked via CARD-CARD interactions, forming a more compact phenotype [Dick et al., 2016, Franklin et al., 2018]; however, it is not clear whether these processes are in fact temporally separated. Another possibility would be that ASC filament extension via PYD-PYD interactions occur concurrently with CARD-CARD interactions, which may either allow two growing filaments to crosslink or allow a monomer to join with an existing filament, forming a new 'branch'. Recent studies have suggested that full-length ASC forms filaments through a combination of CARD-CARD and PYD-PYD interactions, with filaments stacking laterally [de Alba, 2019, Nambayan et al., 2019]; these contrast to earlier studies in which ASC CARD and PYD domains were studied in isolation [Vajjhala et al., 2014, Lu et al., 2014, Sborgi et al., 2015].

There are clearly regulatory systems which limit ASC aggregation, since in endogenous systems, ASC monomers do not appear to interact in more than a transient fashion in the absence of PAMPs/DAMPs. Spontaneous ASC speck formation has been observed in some ASC overexpression systems, but this may be artefactual [Kuri et al., 2017]. It is therefore unclear what mechanisms prevent the clustering of ASC in resting cells. Various regulatory

mechanisms have been proposed, for example, the action of COPs and POPs; similarly to their regulation of NLRs, these may bind to the CARD and PYD domains of ASC and caspase-1, limiting the potential for ASC interaction with NLRs and caspase-1 processing [Humke et al., 2000, Lamkanfi et al., 2004, Lee et al., 2001, Druilhe et al., 2001, Bedoya et al., 2007]. Another regulatory system is negative regulation by IκB kinase α (IKKα), which controls ASC aggregation in the nuclear and perinuclear area; ASC is moved out of the nucleus through the action of IKKi, and the kinase activity of IKKα is limited through the recruitment of PP2A following the introduction of a DAMP/PAMP [Martin et al., 2014]. It has also been shown that various isoforms of ASC exist, as the product of differential splicing; one such isoform may in fact inhibit IL-1β processing [Matsushita et al., 2009, Bryan et al., 2010]. However, as with NLR regulatory processes, it is still unclear how these regulatory processes are themselves controlled, and to what extent they are important.

Interestingly, the behaviour of ASC has been likened to that of prions, and in particular, it has been demonstrated that ASC can recapitulate the behaviour of prion proteins in yeast [Cai et al., 2014]. This suggests that ASC, like NLRs, may normally exist in an unreactive form; however, its propensity for reacting with other ASC monomers may increase upon interaction with an active NLR, and may propagate this increased propensity to further ASC monomers.

1.3.3 Cell-level properties

At the cell level, the inflammasome is a very unusual structure. It contains thousands to tens of thousands of individual proteins (the exact number has not been quantified), and yet it can form within minutes of the introduction of a stimulus. In fact, ASC is depleted from across the cell and absorbed into a single complex [Fernandes-Alnemri et al., 2007]. Multiple ASC specks within a single cell are not often observed, with the exception of some ASC overexpression systems, in which occasionally two, or infrequently, more than two inflammasomes are seen [Man et al., 2014b]. This process is extremely energetically unfavourable, and ASC speck formation is not necessary for pyroptosis or IL-1β production in NLRC4 inflammasome systems, although disruption of speck formation decreases the efficiency of IL-1β production [Proell et al., 2013, Dick et al., 2016]. Moreover, this appears to be a 'digital' response; the output of IL-1β, for example, shows 'all-or-nothing' behaviour [Liu et al., 2014].

Another intriguing aspect of inflammasome formation is the differences between cells

even within the same sample. Even cells treated simultaneously with inflammatory stimuli will form inflammasomes at different times; some will not form inflammasomes at all [Man et al., 2014b]. The latter effect is considerably more marked in NLRP3 systems compared to NLRC4 systems. Of course, some variation should be expected in such a complex biological system, but it is unknown what the primary cause for this variation is, and whether there is a biological benefit which arises from it. Similarly, it is natural to ask whether we can explain the differences between variation within NLRP3 samples and NLRC4 samples due to some difference in the abundance or behaviour of NLRP3 and NLRC4.

1.3.4 Crossing scales

Conceptual models of inflammasome formation often assume a radial or tube-like structure. Radial structures often feature concentric rings of NLRs, with ASC forming a ring inside an NLR oligomer, with caspase-1 gathering in the centre of the structure [Lechtenberg et al., 2014, Brewer et al., 2019]. Alternatively, a ring-like NLR oligomer is portrayed as the seed for a single extended ASC tube-like structure, with caspase-1 attaching alongside or to the ends of this tube [Lu et al., 2014, Shen et al., 2019]. Some slightly more sophisticated inflammasome structure models include multiple ASC filaments which may be cross-linked [Franklin et al., 2018, Broz and Dixit, 2016].

However, since attempts to ascertain the structure of the full inflammasome have been limited to date, these models are based on a number of assumptions. It is unclear how many NLR oligomers are involved in a single inflammasome, and, if these oligomers form extended filaments rather than discs, how large these oligomers are. The spatial arrangement of ASC filaments is also uncertain. In fact, while these radial and tube-like models of inflammasome formation are consistent with established structural information about the domain interactions driving inflammasome formation (and are amenable to representation in two-dimensional cartoon form), they are at odds with established inflammasome imaging. In particular, confocal fluorescence microscopy carried out by Man *et al.*, in which NLRC4, NLRP3, ASC and caspase-1 were labelled in endogeneous cells, demonstrated that a two-dimensional cross-section of the inflammasome shows concentric rings of ASC, NLRC4 and caspase-1 [Man et al., 2014b]. Similarly, imaging by Sanders *et al.* suggests that ASC forms spherical structures with caspase-1 gathering on its outer surface [Sanders et al., 2015]. While it is not clear why these two studies show different arrangements of proteins within

the inflammasome structure, and it is possible that some spatial artefacts may have occurred due to labelling, this view of the inflammasome certainly suggests that several NLR oligomers are involved in a single speck, and that there is a more complex organisational structure than the simple radial or tube model. This view is corroborated by more recent correlative light and electron microscopy of the ASC speck in zebrafish larvae, which show a highly branched structure of ASC filaments [Kuri et al., 2017]; however, this does not give any insight into the interaction between ASC and other inflammasome constituent proteins. It is worth bearing in mind that all of these studies consider only static images, and show only the endpoint of inflammasome formation, so we can infer little about intermediate formation stages.

1.4 Mathematical modelling of the inflammasome

As far as I have been able to ascertain, there has only been one previous attempt to create a mathematical or computational model of ASC speck formation [Cheng et al., 2010]. In this study, Cheng and colleagues carried out live imaging of HeLa cells transfected with yellow fluorescent protein (YFP)-labelled ASC using confocal laser scanning microscopy. The intensity of YFP measured at sampled points within the area of a cell was used as a proxy for the concentration of ASC at that point in the cell; taking intensities over all timepoints in the time series gives a view of ASC depletion over time. Comparing results for points in the cytosol, nucleus and speck of each cell showed a decrease in ASC concentrations in the nucleus and cytosol, while intensities in the speck increased. Furthermore, the distance from the centre of the ASC speck to each sampled point was measured, and the rates of ASC depletion compared to the distance from the speck. It was shown that there was no link between distance from the speck and the rate of depletion of ASC.

The authors then compared the results to a simple model of ASC aggregation. This tracked ASC concentrations in a two-dimensional representation of the cell, including the following reactions between ASC monomers and a single incipient 'speck' complex:

$$\text{free ASC} + (\text{speck})_n \underset{k_{on}}{\overset{k_{off}}{\rightleftharpoons}} (\text{speck})_{n+1} \tag{1.1}$$

Free ASC monomers associate with a growing speck of size n to create a speck of size $n+1$ at a constant rate k_{on}, and ASC monomers dissociate from the speck at a constant rate k_{off}.

The ASC was allowed to diffuse freely through the cell, and the concentration of ASC throughout the domain over time was simulated. The data extracted from the time series were then compared to simulation results when active transport was included, and excluded from, the model. The closest similarities arose for the model with diffusion only, suggesting that active transport is not a factor in ASC aggregation, at least in the transfection system considered here.

The use of fluorescence microscopy as a proxy for ASC abundance is a useful way of extracting quantitative and spatial data describing the distribution of ASC throughout the cell through the process of ASC speck formation. However, there were a number of limitations to both the analysis pipeline used, and the model presented in this study. Firstly, the system of transfecting HeLa cells is somewhat artificial; aggregation of ASC was not initiated by infection or other inflammasome formation-initiating treatment, but rather prompted only by transfection. The resulting data cannot then be directly compared to ASC aggregation explicitly caused by inflammasome formation. The addition of the large YFP tag to the transfected ASC may also have affected the ASC aggregation kinetics. It is unclear how the authors chose sample points from which to extract characteristic intensities across the cell, nor is it clear how the cell was divided into areas corresponding to the nucleus and cytoplasm. Moreover, it is not clearly established that the chosen points are representative of intensities across the compartment in question. Finally, the model of ASC speck formation is extremely simplistic; while it may be representative of aggregation occurring in transfection systems, it assumes that aggregation rate is independent of speck size, and does not capture the different possible modes of interactions between ASC monomers. Other proteins involved in inflammasome formation, such as NLRs, have not been included in the model, and although this would not have been appropriate in the ASC transfection-related aggregation context considered in this study, NLRs may play a much more important role in ASC aggregation in the inflammasome context. Finally, this model assumes that only a single speck is present at any given time, with all other ASC in the cell existing in monomeric form, although no justification is given for this assumption; it would be more natural, especially since no 'seeding' processes have been included in the model, to assume that any ASC monomer can interact with any other ASC monomer.

The absence of a compelling mathematical model of inflammasome formation motivates the creation of a new model which explicitly considers both the other proteins involved in

inflammasome formation beyond only ASC, and which distinguishes between linear growth of ASC filaments and potential crosslinking or branching of these filaments.

1.5 The mathematics of coagulation processes

Despite the lack of mathematical models explicitly depicting inflammasome formation, there is a much broader class of existing mathematical and physical models which have been used to analyse the kinetics of coagulation or aggregation processes. Some popular approaches include a physical perspective, including the study of the formation of micelles, i.e. self-assembled aggregates of surfectants; the modelling of micelle formation was pioneered by Murray and Hartley [Murray and Hartley, 1935]. Another popular approach is the study of the fractal structure of randomly-formed aggregates, which was particularly prevalent in the 1980s; this area focuses on how self-similar structures can arise via diffusion-limited aggregation of particles, and was pioneered by Witten and Sander in 1981 [Witten and Sander, 1981]; a useful review is given by Jullien [Jullien, 1987].

However, in this research we will focus on the simplest possible mathematical models of aggregation. These models often use a mass action formulation (i.e. the assumption that the rate at which a reaction occurs is proportional to the concentration of the reactants). The reactions may be expressed using a deterministic or stochastic approach.

The deterministic approach uses ordinary differential equations (ODEs), or, in the case of spatial models, partial differential equations (PDEs). These equations track the concentration of various reactants over time; in the case of coagulation modelling, these would take the form of $x_i(t)$, the concentrations of clusters of a size i over time. Deterministic approaches are often a useful first recourse for reaction modelling, due to the range of accessible mathematical tools for analysis. In some cases, it may not be possible to solve these equations explicitly, to give the concentration of each reactant over time, especially for more complex, realistic models. Alternatively, it may be useful to use the ODE system to derive the moments $\sum_i i^n x_i(t)$, which can be used to derive properties of the distribution of cluster sizes at time t. The ODE system can also be manipulated to give a PDE for the generating function $G(s,t) = \sum_i x_i(t)s^i$; the generating function uniquely defines the distribution of cluster sizes for each time, and may be used to infer the moments of this distribution. In some cases, it may be possible to use G to derive solutions for x_i. However, it is also pos-

sible to solve ODE systems equations numerically, using a computational integration solver. While this is generally less challenging than finding an analytical solution, the results will only be an approximation of an analytical solution, and the solver may fail to converge altogether for ODE systems with particular undesirable qualities (such as 'stiffness'). Moreover, numerical solutions are only valid for a single parameter set; to understand how changing parameter values affect results, multiple numerical solutions must be calculated. Finally, an analytical approach may be used to consider more global properties of solutions to an ODE system without explicitly deriving these solutions fully; for example, proving the existence and uniqueness of a solution given certain parameters and intial or boundary conditions, the existence and stability of equilibria in the system, and the asymptotic behaviour of the system (in practice, this often means the behaviour of the system on very short or long timescales). However, ODEs often fail to show behaviour which is important in biological systems. Due to their deterministic nature, ODEs are unable to capture the stochastic variation which arises even within the same biological samples. They also fail to capture accurately the dynamics of systems which involve switching or small numbers of reactants, due to the use of continuous concentrations, rather than absolute numbers of reactants, as variables in these systems.

Alternatively, a stochastic approach may be used; the reactions in a system are summarised by a set of propensity functions, which are an analogue of the reaction rates in deterministic systems. The abundances of reactants are tracked as discrete counts over time. These systems may be analysed using a master equation approach; the propensity functions can be used to construct an ODE for the time evolution of $P(x,t)$, the probability that the system is in a state x at a time t. In the case of a coagulation system, x will often take the form of a vector $\boldsymbol{x} = (x_1, x_2, \cdots, x_N)$; in this case, $P(\boldsymbol{x},t)$ is the probability that for $i = 1, \cdots, N$ there are x_i clusters of size i at time t. In some simple cases, it may be possible to solve the ODE directly for $P(\boldsymbol{x},t)$; however, it is much more common to use the master equation to derive equations for the moments, or derive a PDE for the generating function $G(s_1, s_2, \cdots, s_N, t) = \sum_{\boldsymbol{x}} P(\boldsymbol{x},t) s_1^{x_1} s_2^{x_2}, \cdots, s_N^{x_N}$, where the sum is taken over all possible values of $\boldsymbol{x}$. Once again, the generating function can be used to derive the moments of the system, or in some cases, $P(\boldsymbol{x},t)$. Another common approach is to view the system as a continuous-time Markov chain (CTMC); this is particularly useful for analysing the distributions of times at which the system enters an absorbing state (for example, the time at which a cluster of a given size first emerges). Alternatively, the system may be simulated algorith-

mically using a kinetic Monte Carlo method; the Gillespie algorithm is a popular foundation for approaches to simulating reaction models. While technically approachable, these simulation processes are generally considerably more computationally intensive than numerical schemes for integrating ODEs.

It is generally assumed that ODE models and their stochastic analogues will behave similarly, in that ODE models depict the behaviour of the stochastic system in the 'mean-field' limit, as the number of monomers in the system becomes very large. However, for many of the models described below this has not been rigorously proved, and moreover the removal of effects due to small monomer numbers may not be desirable.

One of the key problems that arises with either approach in aggregation systems is the extremely large number of variables involved; assuming that cluster sizes are treated as discrete, each possible size gives rise to a new variable. Indeed, many models of this type do not limit the possible size of clusters, leading to infinitely large ODE or stochastic systems which cannot be closed. The introduction of a generating function can circumvent this issue, as it is a single function which can summarise abundance of all possible cluster sizes.

Another important factor to consider is the initial conditions of the system. There is often an assumption of monodispersity of initial conditions (ICs); that is, that initially all clusters in the system have the same size. In practice, this is generally taken to mean that initially only monomers are present. However, the case of polydisperse ICs (i.e., clusters of different sizes are present) may be particularly of interest in systems where we have nucleation processes or some other source of heterogeneity in cluster population. It cannot be assumed that systems with polydisperse ICs behave analogously to those with monodisperse ICs, and in fact, may lead to substantially different solutions in some models. It is therefore important to establish assumptions about ICs in aggregation models.

In the following sections, I will discuss some of the predominant models of aggregation processes, as well as some extensions which are particularly relevant to the problem of inflammasome formation. I will discuss some of the key approaches used in analysing these models, and their relevance to inflammasome modelling.

Since aggregation models are widespread, with a history extending at least back to the early twentieth century, and applied in numerous fields of study, there is a considerable lack of consistency in terminology used in the literature. In general, for ease, I will use the language commonly used in protein aggregation; individual aggregate subunits will be

(a) Becker-Döring (b) Smoluchowski/Marcus-Lushnikov

Figure 1.3: Two basic coagulation processes; clusters grow linearly by addition of monomers, or by coagulation of clusters

referred to as 'monomers', and monomers which have adhered together will be referred to as 'clusters', except in the case of clusters which are explicitly linear, which I refer to as 'filaments'. In general, I will also assume that such clusters grow at rates independent of their size, and I refer to this phenomenon as 'linear growth'.

1.5.1 Basic processes

There are two main approaches which have dominated the study of the kinetics of protein aggregation, namely models in which clusters may grow only via the addition of monomers, and models in which larger clusters may also coagulate [Aldous, 1999]. The most common formulation of these two aggregation schemes are the Becker-Döring model [Becker and Döring, 1935] and the Smoluchowski model [Smoluchowski, 1916] (Figure 1.3). In this section, I will outline the mathematical structure of these models, give a summary of the common approaches to analysing these models, and describe their limitations.

BD models

The Becker-Döring (BD) model was first proposed in a simplified form by Becker and Döring [Becker and Döring, 1935] and extended by Burton [Burton, 1977]. In the BD model, monomers may join with other monomers or with larger clusters, and individual monomers may also dissociate from clusters, but clusters may not interact with one another. This can be summarised using the reactions in Table 1.2.

Here x_i is the abundance of clusters containing i monomers, and a_i and b_i are aggregation and breakage rates respectively, of clusters of size i. The rates a_i and b_i may be functions of i, to reflect different cluster shapes or aggregation and breakage styles, but a case assuming linear growth has a_i and b_i constant (which I will assume to be the case from henceforth).

Reference	Reaction	Propensity function	Explanation
A	$x_1 + x_i \xrightarrow{a_i} x_{i+1}$	$a_i x_1 x_i$ for $i \neq 1$, $a_1 x_1 (x_1 - 1)$ otherwise	Monomer binds to filament of length i to form filament of length $i+1$
B	$x_i \xrightarrow{b_i} x_{i-1} + x_i$	$b_i x_i$	Monomer dissociates from filament of length i to form filament of length $i-1$

Table 1.2: Reactions in the Becker-Döring model

The BD model is especially popular for analysing systems in which aggregates form filamentous structures, since it can be assumed that most growth occurs via linear monomer addition to the filament end. Notably, the BD model has been used in analysing industrial systems [Coveney and Wattis, 1996], and biological processes such as prion and actin polymerisation [Hu and Othmer, 2011, Davis and Sindi, 2016, Holcman, 2017].

The deterministic approach gives the Becker-Döring equations, a set of ODEs describing the time evolution of reactant concentrations:

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = \sum_{i=1}^{\infty} (b x_{i+1} - a x_i x_1) \tag{1.2a}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = \frac{1}{2} a x_1^2 - b x_2 \tag{1.2b}$$

And for $i > 2$:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = a_{i-1} x_{i-1} x_1 - b_i x_i \tag{1.2c}$$

Note the factor of $1/2$ in the equation for $\frac{\mathrm{d}x_2}{\mathrm{d}t}$; in general, monomers may be added to either end of the filament. However, the joining of two monomers is a symmetrical process so the rate at which such reactions occur is halved. Note also that the factor of $1/2$ is not included for the corresponding reaction in the equation for $\frac{\mathrm{d}x_1}{\mathrm{d}t}$, since two monomers are removed via these reactions.

A common approach to analysing this system is to consider the large-time asymptotic behaviour; the system will approach a steady state dependent on parametrisation, and much of the existing analysis of the deterministic system is concerned with determining this steady state, its stability, and the behaviour of the system as it approaches this steady state [Ball et al., 1986, Wattis and King, 1998].

In general, a master equation approach is used to analyse stochastic BD systems; for example, Davis and Sindi [Davis and Sindi, 2016] analyse the following system with constant reaction rates (notation is altered for consistency with Equation 1.2):

$$\begin{aligned}\frac{\mathrm{d}P}{\mathrm{d}t} = &-\frac{1}{2}\underbrace{a x_1(x_1-1)P}_{\text{A1}} - \underbrace{a\sum_{i=2}^{\infty} x_i P}_{\text{A2}} - \underbrace{b\sum_{i=2}^{\infty} P}_{\text{B1}} + \underbrace{\frac{1}{2}a(x_1+2)(x_1+1)W_1^+ P}_{\text{A3}} \\ &\underbrace{+a(x_1+1)\sum_{i=2}^{\infty}(x_i+1)W_i^+ P}_{\text{A4}} + \underbrace{b\sum_{i=2}^{\infty}(x_{i+1}+1)W_i^- P}_{\text{B2}}\end{aligned} \tag{1.3}$$

where $\boldsymbol{x}(t)$ is a vector in which $x_i(t)$ is the number of clusters containing i monomers present at time t, $P(\boldsymbol{x},t)$ is the probability that $\boldsymbol{x}$ gives the number of clusters of each size at time t, and we define the operator $W_i^{\pm}P(\boldsymbol{x},t) = P(\boldsymbol{x}\pm(\boldsymbol{e}_1+\boldsymbol{e}_i-\boldsymbol{e}_{i+1}),t)$ where $\{\boldsymbol{e_i}\}$ are the standard basis vectors. Terms A1, A2 and B1 correspond to reactions moving out of state $\boldsymbol{x}$, and terms A3, A4 and B2 correspond to reactions which move the system into state $\boldsymbol{x}$. In particular, terms A1 and A3 account for reactions involving two monomers; the propensity function is slightly different to the analogous rates in the deterministic system (See Table 1.2) to avoid the possibility of a monomer reacting with itself. Note in particular the form of term A3; this term accounts for the system reaching $\boldsymbol{x}$ via the merging of two monomers into a filament of length two. Prior to this step, there were x_1+2 monomers in the system, each of which may react with x_1+1 other monomers, so the rate of this reaction is proportional to $(x_1+2)(x_1+1)$. The factor of $1/2$ in terms A1 and A3 are included to account for symmetry in cluster joining, as in Equation 1.2.

An area of particular interest in stochastic BD modelling is determining the time taken for a cluster of a given size to first form. This is essentially a stopping time problem in analysis of a Markov chain; the derivation of this stopping time can therefore be analysed via numerical or asymptotic techniques [Davis and Sindi, 2016, Hu and Othmer, 2011, Michaels et al., 2016]. Notably, Yvinec *et al.* [Yvinec et al., 2012, Yvinec et al., 2016] derive a full analytic solution for first formation times of a cluster of size n in BD systems using the backwards Kolmogorov equations; however, this solution grows in complexity extremely quickly with n.

In summary, the BD model is a popular view of linear aggregation which is particularly amenable to mathematical analysis, but it is not appropriate to the modelling of many processes due to that very simplicity. While linear growth has been observed in many systems, the limitation to linear growth only is not appropriate in many contexts. Furthermore, many

Reference	Reaction	Propensity function	Explanation
K	$x_i + x_j \xrightarrow{a_i} x_{i+j}$	$K(i,j)x_ix_j$ for $i \neq j$, $K(i,i)x_i(x_i-1)$ otherwise	Clusters of sizes i and j merge to give cluster of size $i+j$

Table 1.3: Reactions in the Smoluchowski model

of the features of interest in this system derive from the interplay between growth and fragmentation, which is less relevant in systems such as inflammasome formation.

However, the BD model is the basis of many more complex models of coagulation, notably the nucleated polymerisation model of prion-like protein aggregation, which will be discussed further in Section 1.5.2.

Coagulation models

Coagulation models present a different view of aggregation to that of BD models, allowing merging of clusters of any size. These models have also been used in modelling a diverse range of systems including the formation of aerosols [Drake, 1972, Gillespie, 1972, Ramabhadran et al., 1976, Pruppacher and Klett, 1978], polymers [Stockmayer, 1943], and even celestial bodies [Lee, 2000] and social groups [Gueron and Levin, 1995].

The reactions in coagulation models are summarised in Table 1.3.

Here, the rate of reaction $K(i,j)$ is often described as a 'kernel', and may depend on the size i and j of the reactants. The precise form of the kernel may be as simple as a constant, or a linear combination of $i+j$ and ij. More sophisticated kernels may be used to explicitly take into account the motion of the complexes in space, or the shape of the complex. Summmaries of some common kernels and their uses are given in reviews by Aldous, and Jeldres *et al.* [Aldous, 1999, Jeldres et al., 2018].

Note that it is rare to include breaking or fragmentation in these models, since this generally leads to the models becoming analytically intractable.

The most widespread basis for coagulation modelling is a deterministic approach pioneered by Smoluchowski [Smoluchowski, 1916]. The equations describing the evolution of the number $x_i(t)$ of clusters of discrete size i at time t are as follows:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{1}{2}\sum_{j=1}^{i-1} K(j,i-j)x_jx_{i-j} - x_i\sum_{j=1}^{\infty} K(i,j)x_j \tag{1.4}$$

Note the factor of $1/2$ in the first term, which is included to avoid double-counting of the same complex due to symmetry of i and j. If i is even, there is a central term for which $j = i - j$; however, the factor of $1/2$ is still included, since this corresponds to a reaction in which two clusters of equal size merge. Similarly to Equations 1.2 and 1.3, we include the factor of $1/2$ to account for symmetry.

Alternatively, cluster size can be regarded as continuous, which is justifiable for large populations of reactants over longer timescales. This yields integro-differential equations which may be more amenable to some forms of analysis. Now writing $x(i,t)$ for the number of clusters of continuous size i at time t, Equation 1.4 becomes:

$$\frac{\partial x(i,t)}{\partial t} = \frac{1}{2}\int_0^i K(j,i-j)x(j,t)x(i-j,t)dj - x(i,t)\int_0^\infty K(i,j)x(j,t)dj \tag{1.5}$$

A notable property of the Smoluchowski equations is that for some kernels, the total mass of the system, $\sum_{i=1}^{\infty} ix_i(t)$ in the discrete case, and $\int_0^\infty ix(i,t)di$ in the continous case, is not conserved. In particular, the total mass is constant until a critical time, after which it begins to decrease. Although from Equations 1.4 and 1.5 it would appear that mass is conserved, following the argument of Wattis *et al.*, consider the flux J_r from clusters of size less than or equal to r to clusters of size greater than r [Wattis, 2006a]:

$$J_0 = 0 \tag{1.6a}$$

$$J_r = \sum_{i=1}^{r}\sum_{j=r+1-i}^{\infty} iK(i,j)x_i x_j \tag{1.6b}$$

From the definition of J_r, we also have

$$\frac{\mathrm{d}(rx_r)}{\mathrm{d}t} = J_{r-1} - J_r \tag{1.7}$$

Now summing Equation 1.7 over r gives:

$$\frac{\mathrm{d}\sum_{r=1}^{\infty} rx_r}{\mathrm{d}t} = -\lim_{r\to\infty} J_r \tag{1.8}$$

For some choices of kernel, this limit may be nonzero; consider, for example, a case in which the cluster size distribution decays algebraically for t sufficiently large ($x_i \; i^{-\gamma}$ for some γ as $i \to \infty$); then choosing $K(i.j)$ as a function of i^γ and j^γ leads to a nonzero limit for J_r. There is therefore a flux of mass out of the system due to the formation of a cluster of infinite size known as a 'gel', to which mass is lost; this phenomenon is known as 'gelation', and the critical time at which this cluster emerges is known as the 'gelation time'. The gelation phenomenon was first characterised in the Smoluchowski equations by Van

Dongen [Van Dongen, 1987]. This has often been used as a model for phase transitions in physical systems.

Explicit solutions to the Smoluchowski equations have been calculated in some cases, generally only up to gelation times; notably, Davies *et al.* [Davies et al., 1999] use a generating function to derive solutions to Equation 1.4 for monodisperse ICs and constant kernels, as well as a product kernel of the form $K(i,j) = ij$; substituting the function $G(z,t) = \sum_{k=1}^{\infty} x_k(t)e^{-kz}$ into Equation 1.4 gives a partial differential equation which can be solved using the method of characteristics. The distribution of particle sizes as a function of time can be retrieved by expanding the solution for G in powers of e^{-z}. For example, for a constant kernel $K(i,j) = K$, substituting into Equation 1.4 gives the following:

$$\frac{\partial G}{\partial t} = K\left(\frac{1}{2}\sum_{k=1}^{\infty}\sum_{j=1}^{k-1} x_j x_{k-j} e^{-zj} - \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} x_j x_k e^{-zj}\right) \tag{1.9}$$

Now exchanging the order of summation in the first term and shifting the sum over k we have

$$\frac{\partial G}{\partial t} = K\left(\frac{1}{2}G^2 - GX\right) \tag{1.10}$$

where $X(t) = \sum_{j=1}^{\infty} x_j$. Supposing an initial monomer concentration of 1, the initial conditions are $G(z,0) = e^{-z}$.

An expression for $X(t)$ can be calculated by substituting X into Equation 1.4:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = -\frac{1}{2}X^2 \tag{1.11}$$

This can be solved to give $X = \frac{2}{t+2}$.

Equation 1.10 can now be solved to give the following:

$$G = \frac{4}{(t+2)((t+2)e^z - t)} \tag{1.12}$$

Expanding Equation 1.12 in powers of e^{-z}, $x_i(t)$ is the coefficient of e^{-iz}:

$$x_i = \frac{4t^{j-1}}{(t+2)^{j+1}} \tag{1.13}$$

A discrete stochastic approach analogous to the Smoluchowski equations, known as a Marcus-Lushnikov (ML) process, was proposed first by Marcus [Marcus, 1968]. While the deterministic Smoluchowski equations are expected to give the limit of the equivalent ML process as the number of reactants tends to infinity, in many cases this has not been rigorously established.

These models involve reactions of discrete numbers of reactants in a finite volume; using notation defined for the BD model, we have the following master equation:

$$\begin{aligned}\frac{\mathrm{d}P}{\mathrm{d}t} = &\underbrace{-\sum_{i=1}^{N}\sum_{j=1}^{N} K(i,j)x_i x_j P}_{\text{K1}} + \underbrace{\sum_{i=1}^{N} K(i,i)x_i P}_{\text{K2}} \\ &+ \underbrace{\sum_{i=1}^{N}\sum_{j=1}^{n} K(i,j)(x_i+1)(x_j+1)W_{ij}^{+}P}_{\text{K3}} - \underbrace{\sum_{i=1}^{N} K(i,i)(x_i+1)W_{ii}^{+}P}_{\text{K4}}\end{aligned} \tag{1.14}$$

Here, N is the total number of monomers and $W_{ij}^{+}P(t,\boldsymbol{x}) = P(t,\boldsymbol{x}+(\boldsymbol{e}_i+\boldsymbol{e}_j-\boldsymbol{e}_{i+j}))$, where $\{\boldsymbol{e_i}\}$ are the standard basis vectors. Term K1 accounts for reactions which take the system out of state $\boldsymbol{x}$, and term K3 accounts for reactions bringing the system into state $\boldsymbol{x}$. Terms K2 and K4 are correction terms for terms K1 and K3 respectively, which account for the impossibility of a cluster joining to itself.

Exact solutions for cluster size distribution have been derived for a number of ML processes characterised by ICs and kernel. Notably, Lushnikov has derived exact solutions for linear and product kernels and monodisperse initial conditions using a generating function approach [Lushnikov, 1978]. Lushnikov's approach was extended by Fronczak *et al.*, who provide a generating function-derived solution for arbitrary initial conditions and product kernel [Fronczak et al., 2019].

A particularly interesting approach is presented by D'Orsogna *et al.* [D'Orsogna et al., 2015]. Instead of seeking a solution for the full system, the authors of this study truncate the system and, applying techniques developed in earlier studies of BD systems [Yvinec et al., 2012, Yvinec et al., 2016], use backward Kolmogorov equations to derive distributions of 'first-passage' times for when clusters of given sizes are first formed. A similar approach has been used by Holcman and Hozé in examining the time taken for viral capsids to assemble [Hozé and Holcman, 2014, Holcman, 2017].

In summary, the Smoluchowski and ML models are analogous deterministic and stochastic approaches to modelling coagulation. While fragmentation of clusters is often neglected in these models due to increased analytical complexity, they also offer the possibility of modelling the growth of aggregates through other means than simple linear filament formation. In addition, since dissociation of inflammasomes is not generally observed in endogenous systems, the effects of fragmentation may well be less important in ASC coagulation; it is

certainly less important than capturing the potential for more complex coagulation than only linear growth. Furthermore, although it is tempting to compare gelation to the sudden emergence of a large complex in inflammasome formation, the mass loss that occurs during this transition may be undesirable in this system.

While the reactions in a Smoluchowski or ML model alone would be unlikely to be applicable to a biological system, there is a considerable body of literature examining extensions to these models, including models with input or removal of monomers over time, involving several species of reactant within the same complex, and catalytic reactions. I will examine these extensions in more detail in the following section.

1.5.2 Common variants

NP models

One particularly well-developed instance of coagulation modelling is prions and prion-like proteins. This is an important area of study due to its therapeutic implications; prions and prion-like proteins such as amyloid-β and tau proteins are implicated in various neurological conditions including scrapie, spongiform encephalopathies and Creutzfeld-Jakob disease, and various forms of dementia. However, these proteins also display some behaviours which are both biologically and mathematically interesting. Prion-like proteins are present endogenously in cells in non-pathogenic monomeric forms, but the prevailing biological view is that they may also exist in misfolded pathogenic multimeric forms; monomers may be converted into the pathogenic form via interaction with a pathogenic multimer. These multimers often take the form of filamentous structures (which may themselves aggregate to form larger fibrillar structures or 'plaques'); filaments may also fragment and thus propagate the pathogenic monomer form. Another common assumption is that filaments are unstable below a certain critical size n_c; prion proliferation is therefore 'seeded' by the appearance of one or more nucleating filaments of size n_c. These seeds may appear spontaneously given a sufficiently long timeframe, or may be introduced by an external source. This process is often referred to as 'nucleated polymerisation' (NP).

NP models are of particular interest in inflammasome modelling, due to the prion-like characterisation of inflammasome proteins; in particular, the proposed nucleation and autocatalytic properties of ASC and various NLRs, and the biological link between ASC and

Reference	Reaction	Explanation
A	$x_1 + x_i \xrightarrow{a_i} x_{i+1}$	Monomer binds to filament of length i to form filament of length $i+1$
B	$x_i \xrightarrow{b_i} x_{i-1} + x_i$	Monomer dissociates from filament of length i to form filament of length $i-1$
N	$n_c\hat{x} \xrightarrow{k_n} x_{n_c}$	n_c monomers nucleate to form a single cluster of size n_c

Table 1.4: Reactions in the Oosawa-Kasai NP model

prions [Cai et al., 2014], suggest a model incorporating some of the elements often used in NP modelling.

The earliest NP-type models were developed by Oosawa and Kasai in 1962 [Oosawa and Kasai, 1962] . They present a deterministic ODE model of filament formation with a rate limiting step in which nuclei of size n_c are formed. These nuclei extend into filamentous forms via linear monomer addition; monomers may also dissociate from the ends of filaments. The reactions involved are detailed in Table 1.4 and Figure 1.4a.

Here $\hat{x}$ is the concentration of (non-pathogenic) monomers in the system and x_i is the concentration of filaments of length i, for $i \geq n_c$. The parameters a, b and k_n respectively describe the rates of monomer addition to and dissociation from filaments, and nucleation; it is assumed that these rates are constant. Note that in contrast to the Becker-Döring and Smoluchowski systems, monomers are denoted $\hat{x}$ rather than x_1 to reflect the fact that the monomer population is distinct from the cluster population, and that (with the exception of nucleation) monomers cannot form clusters without interacting with other clusters.

The equations which describe these reactions are as follows:

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = \underbrace{-k_n\hat{x}^{n_c}}_{\text{N}} \underbrace{-a\sum_{i=n_c}^{\infty} x_i}_{\text{A}} \underbrace{+b\sum_{i=n_c}^{\infty} x_i}_{\text{B}} \tag{1.15a}$$

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \underbrace{a\hat{x}(x_{i-1} - x_i)}_{\text{A}} \underbrace{+b\hat{x}(x_{i+1} - x_i)}_{\text{B}} \underbrace{+\delta_{in_c}k_n\hat{x}^{n_c}}_{\text{N}} \tag{1.15b}$$

The Oosawa-Kasai approach was further developed by Masel, Nowak and others [Nowak et al., 1998, Masel et al., 1999]. They proposed an ODE model of prion proliferation including monomer production and degradation, filament clearance, filament growth through monomer

Reference	Reaction	Explanation
C	$\emptyset \xrightarrow{c} \hat{x}$	Monomer created
D1	$\hat{x} \xrightarrow{d_1} \emptyset$	Monomer degraded
D2	$x_i \xrightarrow{d_2} \emptyset$	Filament of length i degraded
A	$\hat{x} + x_i \xrightarrow{a} x_{i+1}$	Monomer is added to filament of length i to form filament of length $i+1$
B	$x_i \xrightarrow{b} x_j + x_{i-j}$	Filament of length i fragments into filaments of length j and $i-j$ where $j \geq n_c$ and $i-j \geq n_c$
	$x_i \xrightarrow{b} x_j + (i-j)\hat{x}$	Filament of length i fragments into filament of length j and $i-j$ monomers, where $j \geq n_c$ and $i-j < n_c$

Table 1.5: Reactions in the Masel-Nowak NP model

addition and filament fragmentation. In contrast to the Oosawa-Kasai model, nucleation via reaction of monomers is not included in the model; filaments arise either through initial conditions or through fragmentation of larger filaments. It is assumed that filaments smaller than the critical size n_c are automatically destroyed, dissociating into monomer subunits. A summary of the reactions involved is given in Table 1.5 and Figure 1.4b.

Here $\hat{x}$ is the concentration of monomers and x_i is the concentration of filaments of length i. The parameters c, d_1 and d_2 respectively describe the rates of monomer production, monomer degradation and filament degradation; a and b are the rates of monomer addition to filaments and filament breakage, either into two smaller filaments, which may immediately degrade into monomers if smaller than n_c. It is assumed that all parameters are constants. Note that we assume that when a filament fragments, at least one of the remaining fragments is greater in size than n_c.

The equations are as follows:

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = \underbrace{c}_{\text{C}} \underbrace{-d_1\hat{x}}_{\text{D1}} \underbrace{-a\hat{x}\sum_{i=1}^{\infty} x_i}_{\text{A}} + \underbrace{2b\sum_{i=1}^{n_c-1}\sum_{j=i+1}^{\infty} i x_j}_{\text{B1}} \tag{1.16a}$$

For $i < n_c$:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = 0 \tag{1.16b}$$

For $i \geq n_c$:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \underbrace{a\hat{x}(x_{i-1} - x_i)}_{\text{A}} \underbrace{-d_2 x_i}_{\text{D2}} \underbrace{-b(i-1)x_i}_{\text{B2}} + \underbrace{2b\sum_{j=i+1}^{\infty} x_j}_{\text{B3}} \tag{1.16c}$$

Note the form of the terms describing breakage, B1, B2 and B3. A filament of length i can fragment in $i-1$ locations, hence the form of term B2. Up to $n_c - 1$ monomers may be created from a cluster of size greater than the number of monomers released, hence the form of term B1; any cluster of size greater than j may fragment to give a cluster of size j, hence the form of B3. The factor of 2 in terms B1 and B3 account for the fact that the monomers may be released from either end of the filament.

Some notable results from Masel demonstrate that this many-dimensional ODE system can be reduced via summation to give a closed system in terms of the zeroth and first moments $X = \sum_{i=1}^{\infty} x_i$ and $Y = \sum_{i=1}^{\infty} i x_i$:

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = c - d_1\hat{x} - a\hat{x}X + bn_c(n_c - 1)X \tag{1.17a}$$

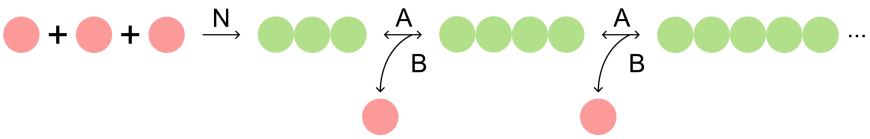
$$\frac{\mathrm{d}X}{\mathrm{d}t} = -d_2X + bY - 2b(n_c - 1)X \tag{1.17b}$$

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = a\hat{x}X - d_2Y - bn_c(n_c - 1)X \tag{1.17c}$$

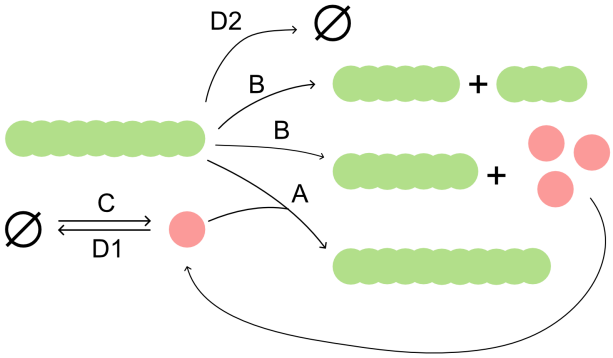
Although this reduced system cannot generally be solved analytically, it is considerably more amenable to analysis than the full system above. Masel also showed that there are clear parallels between this model and an epidemiological model, deriving an 'R_0'-like value which gives critical conditions for expansion of the prion population, and deriving estimates for this parameter from experimental data. A similar model reduction approach can be used with the Oosawa-Kasai model, which can be solved analytically in some cases [Michaels et al., 2017].

Much of the existing literature on these models has centred around an analytical approach. Starting with work by Greer *et al.* [Greer et al., 2006], in which the Masel and Nowak model is adjusted to regard filament size as continuous, many subsequent studies have been concerned with establishing the well-posedness of the problem, existence and uniqueness of solutions, existence and stability of equilibrium points in the system, and asymptotic approximation of large-time behaviours [Prüss and Pujo-Menjouet, 2006, Simonett and Walker, 2006, Engler et al., 2006, Calvez et al., 2009, Helal et al., 2013]. While these analyses contribute rigour to the derivation of solutions and determination of conditions when a disease-state equilibrium may arise, they also have relatively little impact on a biological discussion of prion proliferation.

In a similar vein to the somewhat abstract mathematical analysis of the deterministic Masel and Nowak NP model, there have been various attempts to analyse a stochastic ver-



(a) Oosawa-Kasai. N. nucleation, A. monomer addition B. monomer dissociation



(b) Masel-Nowak. D2. filament degradation, B. filament fragmentation, A. monomer addition, C. monomer creation D1. monomer degradation

Figure 1.4: The two most common forms of NP model

sion of the NP model. For example, a stochastic differential equation approach has been used to analyse the NP model, deriving asymptotic solutions for large times and the distribution of first nucleation times [Doumic et al., 2016, Eugène et al., 2016, Robert and Sun, 2019].

Others have taken a more simulation-based approach to stochastic NP modelling. Notably, Proctor *et al.* [Proctor et al., 2012, Proctor et al., 2013] use a Gillespie-like algorithm to simulate an NP process to model amyloid-β and tau protein proliferation in the brain. Since experimental measurements for parameter values in their model were not available, they used parameter sets sampled using a Latin hypercube sampling method from across a feasible domain informed by experimental results. For each simulation they tracked the total number of clusters over time, and in particular, noted the time at which the cluster density exceeded a threshold corresponding to the appearance of disease symptoms; this gave distributions of characteristics comparable to existing experimental amyloid-β and tau protein datasets. They also investigate the effects of various medical interventions by altering parameters of the model and noting the effect on the measured percentages of disease occupancy. A similar approach is taken by Kashchiev [Kashchiev, 2018], who uses stochastic simulations of an Oosawa-Kasai-like model to ascertain a distribution for times at which the

first filament of size n or larger appears.

Due to its links with experimental work, the field of prion-like protein modelling shows how coagulation models can be successfully compared to experimental data. However, in many ways the existing literature is limited in its applicability to inflammasome systems; as with BD systems, the limitation to linear growth is not ideal. Moreover, much of the analysis relies on the assumption of constant coagulation kernels, which is a considerable simplification of a full coagulation model. While nucleation is certainly an important factor in inflammasome formation, this is caused by NLR activation rather than through spontaneous formation of a sufficiently large seed. Finally, the timescales involved in inflammasome formation are much shorter than those involved in prion modelling (seconds rather than years), and many of the processes included in the Masel-Nowak model, such as protein transcription and degradation, are not significant on the much shorter timescales of inflammasome formation. Therefore, the NP model may form a useful basis for modelling some aspects of inflammasome formation, but some adaptation will be required.

1.5.3 Extensions

While extensive analysis has been conducted on models such as the BD, Smoluchowski, ML and NP models, and it has been demonstrated that they are rich in mathematical complexity, they nevertheless fail to fully capture the complexity of biological systems. In particular, the roles of different protein species, different cellular compartments and different coagulation mechanisms are of particular interest when modelling the inflammasome. In this section I will therefore discuss models which combine the basic models discussed previously, or add elements which may be useful in constructing models of the inflammasome.

Combining linear growth and coagulation of clusters

It is unusual for models to combine several of the basic models previously described; however, some studies have made progress in analysing models incorporating multiple coagulation mechanisms. So far, these appear to focus mainly on deterministic models, perhaps due to the complexity and large number of reactions in such systems.

Few models combine BD and Smoluchowski aggregation, but one study by Wattis *et al.* [Wattis, 2006a] establishes some full solutions and asymptotic results for a deterministic system with both BD and Smoluchowski kinetics.

However, models that combine NP dynamics with Smoluchowski-like coagulation of clusters are somewhat more common, since many prion proteins have been observed to form plaques, consisting of multiple filaments which have separately coagulated; however, these are generally concerned with large-time asymptotic and numerical results [Ramabhadran et al., 1976, Li and Roberts, 2009, Schreck and Yuan, 2013, Mobilia et al., 2003].

Multispecies models

Another important consideration in aggregation systems is that in many cases, multiple species of monomer are involved. There have been a number of attempts to extend coagulation models to include multiple species, both determinstic [Lushnikov, 2014, Ke et al., 2003, Fernández-Díaz and Gómez-García, 2010] and stochastic [Trautmann, 1993, Laurenzi et al., 2002, Lushnikov, 2006]. Notably, analytical solutions have been derived for the full deterministic model for a product kernel and arbitrary numbers of species using a Laplace transform approach [Fernández-Díaz and Gómez-García, 2010]. The stochastic system is naturally somewhat more complex, although Lushnikov has derived solutions for the cluster size distributions over time for the two-species model with the kernel $K(i_1, i_2, j_1, j_2) = i_1 j_2 + i_2 j_1$ (where i_l and j_l are the numbers of monomers of species l in the clusters involved in a joining reaction).

However, these models assume a symmetry in the behaviour of monomers of different species; the only difference in their behaviour is encoded in the coagulation kernel. In systems such as inflammasome formation, however, the different species of proteins involved may drive different forms of aggregation; for example, NLRs may have a similar role to 'seeds' in a NP model, but ASC forms filaments, and may or may not branch.

Intracluster reactions

So far there has been little discussion of the morphology of clusters involved in aggregation processes. In many aggregation studies mentioned above, it is assumed that any relevant properties of clusters involved in reactions are encoded in the reaction rates or kernels. However, it is also useful to examine these properties more explicitly. For example, a number of studies have explicitly considered the growth of clusters in Smoluchowski systems which allow include branching [Hendriks and Ernst, 1984, Noureddini and Timm, 1992], intracluster crosslinking of monomers [Galina et al., 2002, Lazzari et al., 2014], or cluster compaction

[Wattis, 2006b]. A model of ovalbumin filament formation also considers a special case of a linear model where end joining in a filament is permitted (so that a closed loop structure is formed) [Kalapothakis et al., 2015]. In general these approaches rely on numerical or asymptotic solutions, with the exception of the study of branching by Noureddini *et al.* [Noureddini and Timm, 1992], and the compaction study by Wattis [Wattis, 2006b], which give exact solutions.

Spatial models

While kernels and reaction rates of aggregation processes may be constructed to take cluster morphology into account, in many cases they are also adjusted to take into account the effects of spatial properties of the system such as diffusion, without explicitly adding a spatial element to the model; for example, a Brownian kernel of the form $K(i,j) = (i^{1/3}+j^{1/3})(i^{-1/3}+j^{-1/3})$ may be used to model systems in which clusters undergo Brownian motion. However, while this kernel-based approach can account for the change in reaction rates due to movement of reactants, it still assumes a well-mixed system; the approach cannot introduce spatial inhomogeneities in initial conditions or monomer source terms, nor can the solutions demonstrate spatial inhomogeneity.

The simplest models to consider space explicitly divide the spatial domain into two compartments. While this is still a fairly unrealistic depiction of the domains in which most aggregation processes take place (especially in biological contexts), it greatly increases the complexity of the system; as well as introducing reactions accounting for movement of monomers and clusters between compartments, the reaction system must involve twice the number of variables (as there must be duplicate variables for each compartment). This has been considered both from a Smoluchowski [Hussain et al., 2014] and a ML [Siegmund-Schultze and Wagner, 2006] perspective. Both of these studies rely on asymptotic and numeric solutions, although Siegmund-Schultze observes that the majority of the variation between actualizations of the system occur in late stages, where there are only a few large clusters present; thus, movement of a single complex from one compartment to another leads to a very large change in the total number of monomers per compartment.

Extending this methodology to more than two compartments suggests the use of a network approach, with nodes corresponding to compartments, and edge weights corresponding to rates of movement between compartments [Shirvani and van Roessel, 2004, Bertsch

et al., 2017, Fornari et al., 2019, Fornari et al., 2020]. Alternatively variables could be taken as functions of continuous space [Herrero et al., 2000, López-López et al., 2005, Olesen et al., 2005, Matthäus, 2006, Krapivsky, 2015, Achdou et al., 2013, Cohen et al., 2014, Bertsch et al., 2017, Fortin, 2019]. In general, analytical solutions are not possible, so these studies have presented numerical, simulation or asymptotic results.

1.5.4 Modelling the inflammasome

An ideal inflammasome model would contain a combination of the models described above; there does not appear to be an existing model that encompasses all of these processes, and it would be extremely complex. We therefore require a minimal model of inflammasome formation which will include the aspects of the system which we are most concerned to investigate, while excluding aspects which would introduce undue complexity. This model should include multiple proteins, which may individually grow in a linear fashion, similarly to the NP models, but the possibility of branching growth and interaction between different species of protein is also included. In addition, it would be useful to incorporate the possibility of clusters merging in a Smoluchowski or ML-like fashion. Including intracluster reactions may be useful for modelling crosslinking within the inflammasome once it is formed. Finally, adding spatial structure to the model, at least separating the cell into nuclear and cytosolic compartments (similar to the model presented in [Cheng et al., 2010]) would allow the differing dynamics in these compartments to be analysed.

Both stochastic and deterministic modelling techniques may be useful here; in many ways, a stochastic depiction of inflammasome modelling would be most natural, especially given that the final state of the system is a single cluster. Stochastic modelling also provides the opportunity to investigate the source of variation in inflammasome formation between technical replicates, which could be compared to different actualizations of the stochastic system (or different runs of a stochastic simulation). However, deterministic modelling is more amenable to analysis, and may still give useful insight into the behaviour of the system.

It is interesting to note that much of the existing coagulation modelling is fairly abstract in nature; with the exception of NP modelling (which is explicitly informed by biological processes) and some industrial applications, many of the studies described above are concerned with mathematical manipulation of the models. While exact solutions are often useful, for many models described above they cannot be obtained, and much of the existing analy-

sis relies on asymptotic results for large times or analysis of gelation processes. Neither of these is of particular concern in inflammasome modelling; in the former case, the monomer pool in inflammasome formation is limited, and the end result of the system not particularly interesting (all reactants are absorbed into a single cluster), and in the latter case, although we are interested in the emergence of a single large complex, this does not correspond to the infinite-mass situation which is described by gelation.

1.6 The biology of lesion formation

1.6.1 Early stages of *Salmonella* infection

Salmonella infections are usually acquired orally, via consumption of contaminated food in the case of humans. Both typhoidal and nontyphoidal serotypes cross the intestinal wall and enter the bloodstream either via endocytosis into M-cells or by passing through tight junctions [Hopkins et al., 2000, Rescigno et al., 2001, Paradis et al., 2021]. Typhoidal serovars may also be internalised by CD18-positive leukocytes and breach the gastrointestinal wall within these cells [Vazquez-Torres et al., 1999].

The bacteria can then pass directly into the bloodstream, or reach the bloodstream via lymph vessels and nodes. Typhoidal serovars travel either extracellularly or inside leukocytes [Vazquez-Torres et al., 1999], and ultimately arrive at the liver and spleen, where they are internalised by resident phagocytes [Richter-Dahlfors et al., 1997, Salcedo et al., 2001].

In general, it is advantageous for the bacteria to reside and replicate within native cells; *Salmonella* has evolved numerous features to evade detection and clearance mechanisms within phagocytes by which they have been engulfed; this is critical to the virulence of *Salmonella* [Fields et al., 1986]. Phagocytic cells rely on reactive oxygen and nitrogen species (ROS and RNS)-dependent processes to limit intracellular bacterial growth; however, these can be subverted by the Type 3 Secretion System (TTSS) in *Salmonella* which inteferes with localization of ROSs and RNSs to the phagosome and the production of ROSs, allowing replication of bacteria within the cellular environment [Mastroeni et al., 2000, Vazquez-Torres et al., 2000, Salcedo et al., 2001, Cirillo et al., 1998, Helaine et al., 2010]. This ability to persist in the intracellular environment allows bacteria to avoid the dangers of the extracellular environment; in particular, extracellular bacteria are vulnerable to opsonisation which increases ROS-dependent antimicrobial function of phagocytes with which

they come into contact [Fields et al., 1986, Rossi et al., 2019]. Eventually, however, the compromised cell is removed via cell death, but this may allow further dispersal of bacteria, promoting the spread of the infection [Lindgren et al., 1996, Brown et al., 2006]. Therefore, it is necessary for the innate immune system to find a balance between cell death processes, which prevent bacteria replicating within the cell, and release of bacteria from lysed cells.

1.6.2 Formation and morphology of legions

As discussed previously, a key characteristic of *Salmonella* infections is the formation of lesions (i.e. a cluster of infected phagocytic cells) in the liver and spleen. For ease, from now on I will consider only the processes of inflammasome formation in the liver, since more data is available to parametrise the model, although the model laid out in later chapters could theoretically also be applied to the spleen with some changes in parametrisation. An overview of stages in lesion formation is given in Figure 1.6.2.

Early in the systemic infection, bacteria reside in individual isolated resident cells in the liver. Infected cells, or cells which have had extensive contact with bacteria, release various cytokines including TNF-α, IL-12, IL-18 and interferon (IFN)-γ. TNF-α acts as a chemokine, attracting other immune cells to the site of an infection, and has a particularly important role in lesion formation, and thus also limitation of disease spread, since lesions cannot form correctly — or at all — in its absence [Mastroeni et al., 1991, Mastroeni et al., 1992, Mastroeni et al., 1993, Everest et al., 1998]. IFN-γ also plays a key role, although this appears to be more associated with bacterial killing and prevention of bacterial infiltration of cells than lesion formation [Kagaya et al., 1989, Muotiala and Mäkelä, 1990, Nauciel and Espinasse-Maes, 1991]. IL-12 and IL-18 promote IFN-γ production by other immune cells [Mastroeni et al., 1996]. IL-12 is also responsible for immunosuppresion during acute infection [Mastroeni et al., 1996]. It has also been demonstrated that IL-4 and IL-15 may be important in clearance of *Salmonella* infection [Everest et al., 1997, Hirose et al., 1999]. There is also a role for cellular adhesion molecules such as ICAM-1; these are expressed in regions near lesions. However, it is unclear how their presence affects lesion formation or bacterial clearance, and experiments comparing the response of naive and vaccinated wild-type and ICAM-1 knockout mice suggest a link with acquired, rather than innate, immune processes [Clare et al., 2003].

Cytokines are also released through pyroptosis since this is a lytic form of cell death,

and the cell contents are dispersed. As more cells are attracted by chemokines and become infected in a localised area surrounded by healthy tissue, a lesion is formed [Richter-Dahlfors et al., 1997].

At early times, lesions are dominated by recruited neutrophils [Richter-Dahlfors et al., 1997, Conlan, 1997]. At later times, around a day into the infection, the neutrophils are replaced by macrophages; lesions are dominated by monocyte-derived macrophages, which arrive via the bloodstream [Hormaeche and Mastroeni, 1990]. It is unclear to what extent Kupffer cells (resident macrophages in the liver responsible for immune surveillance) are involved in this process; these cells derive from an entirely different lineage from the monocyte-derived macrophages, which arrive via the bloodstream [Beattie et al., 2016]. In some cases, resident macrophages have been observed to remain at inflammatory sites, but do not appear to be depleted by the processes of infection and cell death which the monocyte-derived macrophages undergo, although they have been shown to phagocytose *Salmonella* [Nnalue et al., 1992]. The motility of Kupffer cells may be an important factor; the majority of Kupffer cells appear to be sessile, although there is some disagreement on the extent to which Kupffer cells are motile [Frevert et al., 2006].

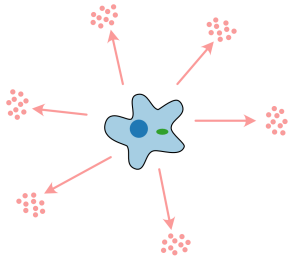
The lesions evolve into spatially separated clusters which are approximately spherical in shape, and around $50 - 100\mu m$ in radius [Richter-Dahlfors et al., 1997, Man et al., 2014a]. Each cell shows low bacterial counts, with the majority of cells containing a maximum of less than ten bacteria (this number is higher in immunocompromised hosts) [Sheppard et al., 2003, Helaine et al., 2010]. The rate of bacterial replication decreases with the bacterial load in that cell, although the lysis rate remains constant [Brown et al., 2006].

Bacteria are released from the lesion when cells within the lesion undergo cell death; in general this is thought to be via pyroptosis, although it has been suggested that apoptotic cell death plays a role in dissemination of bacteria [Richter-Dahlfors et al., 1997, Grant et al., 2008b, Yu et al., 2020]. The bacteria released go on to infect other phagocytes and form further lesions [Sheppard et al., 2003]. Interestingly, lesions appear to be completely functionally independent, bacteria released from one lesion do not spontaneously join other lesions, and the reinfection rate of phagocytes which have already been infected is very low [Sheppard et al., 2003, Gog et al., 2012]. Furthermore, mature lesions do not increase significantly in size; instead, as the total bacterial burden grows, the total number of lesions increases [Sheppard et al., 2003].

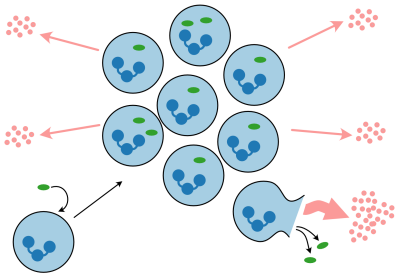
In the delicate balance maintained by the immune system between the removal of cells which have been compromised, and the prevention of bacterial growth, it is unclear what role lesion formation plays. Failure to form lesions leads to rapid growth and spread of bacteria [Conlan and North, 1992], but it is not clear whether this is directly related to the spatial structure of lesions, or whether lesion formation is simply a side effect of the combined dynamics of chemokines, phagocytic cells and bacteria. One interesting theory is that at later times, fibroblasts are attracted to lesions by the chemokine gradient and form a barrier surrounding the lesion, possibly to prevent the escape of bacteria to form further lesions. This theory is analogous to the process of foreign body granuloma formation, a reaction which occurs in response to implanted biomaterials; in particular, neutrophils are first recruited to the implant site, then prompting the recruitment of macrophages, which attempt to phagocytose the implant (thus far, this process is analogous to lesion formation). The recruited macrophages then amalgamate into foreign body giant cells, before fibroblasts are recruited to form a dense outer 'capsule' around the implant [Anderson et al., 2008, Morais et al., 2010, Sheikh et al., 2015, Dondolossa et al., 2016]. It would therefore be useful to investigate whether the recruitment of fibroblasts or another similar process would be necessary for stable lesion formation.

1.6.3 Unanswered questions

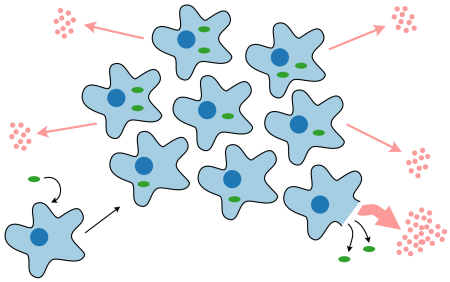
The existing literature on *Salmonella* lesion formation leaves some room for clarity, which could be expanded upon using mathematical modelling. Firstly, it is unclear how the spatial structure of lesions forms at all. A spatial mathematical model would be useful in determining whether chemokine gradients are sufficient for spatial patterns to form, and, in combination with reaction and diffusion processes, remain stable over time, or whether another mechanism (such as the lesions becoming enclosed by a layer of fibroblasts) is necessary to maintain the shape of lesions. Secondly, it is unclear what prevents the spread of bacteria and cells between lesions; it is possible that cellular stiffening and reduction of motility plays an important role here [Man et al., 2014a], but this requires further investigation. Thirdly, it is unclear what role different cell death pathways may play in lesion formation — in particular, to what extent apoptosis plays a role alongside lytic cell death.



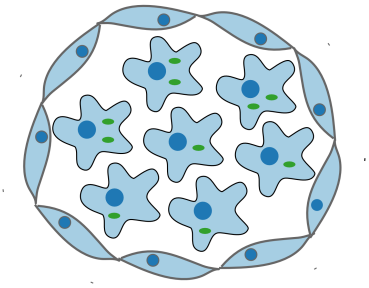
(a) An infected resident phagocyte produces chemokines to attract other phagocytes



(b) Neutrophils are recruited and infected to form a lesion. Neutrophils in the lesion produce chemokines, and lyse, releasing bacteria and a large amount of chemokines.



(c) Neutrophils are replaced by macrophages. Macrophages in the lesion produce cytokines and chemokines, and lyse, releasing bacteria and a large amount of chemokines.



(d) It has been suggested that at later stages, lesions are surrounded by a layer of fibroblasts which prevent bacteria escaping.

Figure 1.5: Stages in *Salmonella* lesion formation

1.7 Mathematical modelling of tissue-level innate immune and inflammatory processes

There have been various previous attempts to model innate immune and inflammatory processes at the tissue or organ level. In general, these models track the abundance of different varieties of cell in a tissue; these often include macrophages, neutrophils, monocytes and resident cells in the tissue. In the case of models of infection, infectious agents such as bacteria or viruses are also tracked. Finally, other relevant substrates such as pro- or anti-inflammatory mediators, medical treatments, or downstream products of the system which are of particular interest (for example, biomarkers of disease) may be included. Throughout this section, I will refer to all entities represented by a variable in this model as 'reactants'.

In the case of spatial models, the abundance of each of these reactants at each location is noted separately (or alternatively, for some classes of model, the location of each individual cell or other reactant may be tracked over time). Reactants may interact only when they are in close spatial proximity.

Most tissue-level models of inflammation can be categorised as one of the following:

- Non-spatial, continuous, deterministic ODE models
- Spatial, continuous, deterministic PDE models
- Spatial, discrete, stochastic agent-based models

In this section, I will document applications of each of these types of model in modelling inflammation in tissues, explaining the mathematical form of the class of model, giving examples of the scenarios which have been modelled using this methodology, and examining in more depth models which have been particularly influential in the development of this field, or which are particularly relevant to the modelling of *Salmonella* infection.

In particular, it is important to note that in developing mathematical models, especially of systems as complex as the innate immune response to infection, there is inevitably a trade-off in determining the level of detail to be included in the model. Simpler models have the benefit of parsimony; reducing the system to its most basic process often gives more insight into the components of the system which are most important in achieving a given outcome. In addition, more complex models may rely on the availability of many parameter values, some of which may not have been measured experimentally; attempting to fit these

parameters to experimental data could easily lead to overfitting. On the other hand, it is important not to oversimplify a model to the extent that key elements of the process are removed. Furthermore, in some contexts, a modeller will be less concerned with illuminating the underlying mechanisms of a system and more concerned with creating a model with more predictive power (for example, for *in silico* experimentation such as tests of therapeutic interventions), in which case, a complex model may be more appropriate.

1.7.1 Nonspatial ODE models

The simplest class of model is the nonspatial ODE model. Nonspatial ODE models assume that the reactants are homogeneously distributed over the spatial domain, and only track the net change in their abundances over the whole domain. The terms in the ODE models correspond to processes which change these abundances, which may be intrinsic to a particular reactant (for example, birth and death of cells and bacteria, or degradation of substrates), or could correspond to the interaction between reactants (for example, killing or phagocytosis of bacteria, production of substrates by cells, or reactions between substrates). As for the aggregation models described above in Section 1.5.1, these might take a simple mass action form (rates of change are directly proportional to the abundance of the reactants involved in the interaction); however, in many cases, it may be more appropriate to use a form of term which reflects saturation at high abundances, particularly in processes involving bacterial or substrate uptake by cells. A common form is a 'Michaelis-Menten-like' sigmoid term. For two arbitrary variables R_1 and R_2, this may take the form $\frac{\mathrm{d}R_1}{\mathrm{d}t} = -R_2 \frac{R_1}{K+R_1}$ for a constant K. This would reflect a situation in which R_2 depletes R_1, and the rate of uptake generally increases with R_1, although the rate reaches a plateau as R_2 becomes saturated. The constant K is half of the value of R_1 at which R_2 becomes saturated.

Such ODE systems can easily be solved numerically, and in some cases analytically. Another common approach is to identify fixed points in the system, and determine their stability; these fixed points give insight into the ultimate fate of the system. Bifurcation analyses can give insight into how changing the parametrisation of the model affects the number, location and stability of fixed points. In the case of inflammation in particular, different configurations of fixed points are often interpreted as different clinical outcomes for the inflammatory process. Bifurcation analysis can thus give insight into how parameter values affect the outcomes of inflammation in different scenarios.

ODE models have been used to investigate inflammation in numerous contexts, notably including wound healing [Nagaraja et al., 2014, Cooper et al., 2015], traumatic brain injury [Vaughan et al., 2018], psoriasis [Valeyev et al., 2010], rheumatoid arthritis [Baker et al., 2013], necrotising enterocolitis [Arciero et al., 2010, Arciero et al., 2013] and influenza infection [Price et al., 2015]. However, the majority of models of inflammation either focus on the behaviour of immune system constituents rather than the specific inflammatory trigger ('generic' models), or describe the scenario of bacterial infection.

The earliest significant ODE model of inflammation was developed by Lauffenburger and Keller as early as 1979 [Lauffenburger and Keller, 1979], and later refined by Alt and Lauffenburger [Alt and Lauffenburger, 1985]; these models were specifically developed to investigate the effects of transport of innate immune cells and bacteria on inflammation, and were used to calculate early estimates of the chemotaxis coefficient and rate of phagocytosis. The model was also used to explain clinical observations of severe infections in patients with defects in leukocyte chemotaxis. The model consists of a very simple three-component ODE system describing interactions between phagocytes, bacteria and an 'attractant' chemokine. The model includes simple linear and mass action terms corresponding to death of bacteria and phagocytes, phagocytosis of bacteria, release of attractant by phagocytes, and birth of bacteria. Significantly, this model indirectly includes chemotaxis of the phagocytes towards the attractant; however, since this is a non-spatial model, the chemotaxis process is defined as introduction of phagocytes from outside the system at a rate proportional to the attractant concentration. An alternative characterisation was later proposed by Lauffenburger and Kennedy [Lauffenburger and Kennedy, 1981]; this model is a two-component ODE system with variables representing bacteria and phagocyte abundance, without explicit inclusion of an 'attractant'. However, this model includes a more sophisticated characterisation of the population dynamics of the reactants, notably including a sigmoid population growth term for bacteria. The presence of two steady states was also demonstrated, which could correspond to a disease-free state, and chronic or acute infection.

Analysis of the above models was mostly confined to considering the steady states of the system, and thus the ultimate fate of the system, although notably different parametrisations may give rise to systems with a single steady state, or a bistable system, in which either an equilibrium or elimination state is reached, depending on ICs.

Other influential early models were developed by Kumar *et al.* [Kumar et al., 2004] and

Reynolds *et al.* [Reynolds et al., 2006]. The former model attempts to explain the dynamics of acute inflammation following bacterial infections using a three-component ODE system including a general bacterial pathogen and early and late proinflammatory mediators. This model is unusual in that it does not explicitly include immune cells. In contrast, the model presented by Reynolds *et al.* includes a bacterial pathogen, neutrophils, an anti-inflammatory mediator and a variable which measures overall damage to the tissue resulting from inflammation. The latter variable is a useful tool to link the inflammatory process directly to clinical outcomes. In both cases, bifurcation analyses were carried out on the respective models, identifying areas in the parameter space corresponding to different possible fates for the system, which in turn correspond to different clinical presentations of sepsis. In particular, the model by Kumar *et al.* identifies possible conditions for persistent or recurrent infections.

A particularly persuasive model was developed by Dunster *et al.* [Dunster et al., 2014]. This is a generic ODE inflammation model including five components, namely apoptotic and active neutrophils, macrophages, and pro- and anti-inflammatory mediators. A series of slightly different models was considered to refine model design; for example, models with and without an anti-inflammatory mediator, and with and without a positive feedback loop in which pro-inflammatory mediators are released by neutrophils. Once again, a bifurcation analysis approach was used to assess the different possible outcomes of inflammation.

More recently, the development of much more complex models has been more common. For example, Shi *et al.* [Shi et al., 2015b] developed a very large ODE system describing bacterial infections of the liver, including Kupffer cells, bacteria, neutrophils, damaged tissue, monocytes, hepatocytes, TNF-α, HMGB-1 (another pro-inflammatory cytokine) and IL-10. This model is unusual in that it includes a large class of specifically-defined cell and cytokine types, specifically including resident cells which are not directly included in the inflammatory process. This gives a useful insight into the wider downstream effects of inflammation; however, in contrast to the models described above, the complexity of this model is such that stability analysis would have been unfeasible, and the authors' analysis was conducted using only numerical solutions to the model. The model was used to reproduce clinical findings regarding the lack of efficacy of anti-TNF-α treatments in resolving sepsis; however, this model arguably suffered from a lack of parsimony in terms of its ability to elucidate the underlying processes.

Finally, a recent model presented by Torres *et al.* exemplifies another approach [Torres

et al., 2019]; this model, of infection in the peritoneal cavity, is an ODE system with components describing neutrophils, active and inactive macrophages, a generic inflammatory stimulus and a pathogen. This study was particularly interesting since it was deliberately designed to fit an *in vivo* experimental dataset, with the aim of developing a tool for hypothesis testing of therapeutic interventions. The authors conducted a very thorough goodness-of-fit and sensitivity analysis for each of the parameters in the model.

There is a clear development of ODE models from simpler models, for which bifurcation analysis is possible, to more complex models, which include many interacting elements but which realistically can only be analysed numerically. Adding complexity to the system comes at a cost of begin unable to use tools such as bifurcation analysis to gain insight, and potentially obfuscates the most important underlying mechanisms driving particular behaviours or clinical outcomes.

A common difficulty is the lack of accurate experimentally-derived parameter values; this is less problematic when carrying out bifurcation analysis, which attempts to take a global view of the parameter space, but is a significant issue for models with a large number of parameters (except in the case of the Torres *et al.* model, which was specifically motivated by an existing experimental dataset).

It is notable that so many of these models omit biologically important elements of the system (for example, the omission of cytokines in the Lauffenburger and Kennedy model, or immune cells in the Kumar *et al.* model). These omissions demonstrate that the reactants which have been removed may not be necessary in creating a system which demonstrates mathematically interesting and relevant behaviour, although we may wish to include them in a model to understand their behaviour. However, it is particularly notable that none of the models described above give a clear account of the dynamics of bacterial spread. Of course, this is unavoidable when using generic inflammation models, but even the models outlined above which explicitly portray bacterial infection do not tend to include phagocytosis of bacteria, reproduction of bacteria within cells, or cell lysis, focusing instead on the behaviour of extracellular bacterial growth and death.

Despite their differences, all of the models described above were created with the ultimate aim of understanding the endpoints of the system, corresponding to long-term outcomes. This is a particular aim of steady-state and bifurcation analysis, which focuses on long-term, equilibrium behaviour of the system. While often especially relevant to clinical

applications, these approaches give less insight into how those long-term states arise. Furthermore, the assumption of spatial homogeneity is a considerable simplification of tissue structure in organisms, and may lead to overlooking interesting spatial behaviour occurring during inflammation (the formation of lesions during *Salmonella* infection is a clear example). This motivates the development of models which do include a spatial element, albeit at the cost of the simplicity and tractability of the ODE models described above.

1.7.2 Spatial PDE models

The simplest generalisation of the non-spatial ODE models are ODE models in which the spatial domain is divided into discrete compartments, with variables in the ODE corresponding to abundances of each reactant in each compartment. This approach has been used to model the spread of infection at the whole-organism level, with compartments corresponding to different organs [Grant et al., 2008a]. However, on the tissue level, there is no such clear delineation into compartments, and a more natural approach when exploring spatial structure is to use spatial PDEs.

Many spatial PDE models are, or are very similar to, reaction diffusion models; these traditionally describe chemical reactions in a spatial setting, and consist of a PDE system with spatially-homogeneous terms describing reactions (which may take the same form as those in the ODE models described above), and diffusion terms of the form $\frac{\partial R}{\partial t} = D\nabla^2 R$ for a variable R, where ∇^2 is the Laplace operator and D is the diffusion coefficient.

It also is common for models of inflammation to include terms for chemotaxis, which is particularly important when modelling the movement of immune cells towards areas with high chemokine concentration. The form usually used to describe chemotaxis in a reaction-diffusion system was originally derived by Keller and Segel [Keller and Segel, 1970, Keller and Segel, 1971]. For a cell abundance R and a chemokine concentration c, a chemotaxis system generally takes the following form:

$$\frac{\partial R}{\partial t} = D\nabla^2 R - \nabla \cdot (\chi R \nabla c) + \text{spatially homogeneous terms} \tag{1.18}$$

Here, χ is a chemotaxis coefficient which in theory may be a function of R and c but in practice is often constant.

In general, these reaction-diffusion systems with or without chemotaxis cannot be solved analytically; numerical solutions may be obtained, although this is generally somewhat more

computationally intensive than for ODE systems. An alternative approach, pioneered by Turing in 1952 [Turing, 1952], is Turing instability analysis. This approach is used to establish necessary conditions on the parameters of a reaction-diffusion model for stable spatial pattern formation of any of the reactants to take place. Essentially, this can occur when a fixed point in a spatially homogeneous version of the system (i.e., the fixed point of the ODE system obtained by removing diffusion and chemotaxis terms) becomes unstable once small spatial perturbations are introduced. This approach can give considerable insight into the long-term behaviour of a reaction-diffusion system without the need for analytical or numerical solutions, and is easily generalised to systems containing chemotaxis terms.

As with ODE models, spatial PDE models have been used extensively in modelling inflammatory processes, including plaque formation in the central nervous system in multiple sclerosis [Lombardo et al., 2017, Moise and Friedman, 2021], foreign body reactions to medical implants [Ibraguimov et al., 2012, Salomonsky and Segal, 2017], necrotising enterocolitis [Barber et al., 2013] and wound healing [Arciero et al., 2011]. However, once again, the most prominent models have been generic inflammation models, or models of bacterial infection. It is worth noting that spatial PDE models are most associated with processes that involve stable spatial structures or patterns, such as plaques or lesions, since these situations are most appropriate for Turing instability analysis.

The earliest influential spatial PDE model of inflammation was developed by Lauffenburger and Kennedy in 1983 as a simple spatial extension of their 1981 ODE model, described above. Their main focus was on deriving stable solutions, and examining the behaviour of the net population sizes. However, this study was primarily concerned with the differences in net outcome when the spatial and analogous non-spatial models were compared, rather than in examining spatial structure formation; furthermore, only a one-dimensional system was considered.

Two models of inflammation particularly relevant to the problem of *Salmonella* infection were developed in 2012; one by Penner *et al.* examining pattern formation in generic skin inflammation [Penner et al., 2012], and one by Pigozzo *et al.* modelling microabcesses formed due to bacterial infection in a generic tissue [Pigozzo et al., 2012]. The Penner *et al.* model is a three-component PDE system including macrophages and a pro- and anti-inflammatory chemokine; stable and travelling wave solutions are derived, as well as different possibilities for spatial pattern formation including spots, stripes and gridlike structures. This model was

used to demonstrate that, provided inhibitor dynamics are sufficiently slow, spatial structures can form in solutions which are comparable to rashes forming on the skin. The Pigozzo *et al.* model is much more complex, with variables representing abundance of living and dead bacteria, neutrophils, active and inactive macrophages, proinflammatory cytokines, and healthy and unhealthy tissue. The model also includes crowding effects, which is unusual for this class of model. The main focus of this study was to develop computational solutions for the system for a single microabcess, tracking its development over time, in order to reproduce experimental measurements of abcess growth. In fact, this model was arguably too complex, since the numerous processes included led to difficulties interpreting the model and, in particular, understanding which processes are most important in determining the ultimate behaviour of the reactants; it is also too complex to be amenable to analytical techniques such as Turing instability-style analysis. The model also includes numerous processes, including production of chemokines by numerous cell types, which had not been parametrised based on experimental data; the parameters in question had to be fitted in order to replicate experimental results, and it is unclear how reliable these estimates are.

Finally, notably, a generic inflammation model was presented recently by Bayani *et al.* [Bayani et al., 2020a]; this model is a five-component spatial PDE system including active and apoptotic neutrophils, macrophages, and pro- and anti-inflammatory mediators. This study includes both a bifurcation analysis of spatially homogeneous solutions and computational solutions revealing possible spatial patterns for results, which establish how aberrant spatial structures can lead to persistent inflammation. However, none of the end states identified in this model correspond to the characteristic lesion pattern seen in *Salmonella* infection.

Overall, although the PDE models described above do give more insight into the spatial behaviour of inflammation systems than ODE models, there are still limitations here. The difficulties of accurate parametrisation has not been averted, and the trade-off between model complexity and amenability to analysis has not been resolved.

Furthermore, although the Penner *et al.*, Pigozzo *et al.*, and Bayani *et al.* models described above could certainly be used to model *Salmonella* infection in the liver, none of these have explicitly included the dynamics of bacterial infection and replication within phagocytes, nor do they differentiate between apoptotic and lytic cell death, or consider the importance of the latter in inflammatory processes . Furthermore, none of these models

demonstrates the characteristic development of multiple lesions throughout the tissue.

1.7.3 Agent-based models

The final common class of model of inflammation is agent-based models (ABMs); in contrast to ODE and PDE models, these are generally discrete and stochastic, and entirely computational simulation-based. In these models, cells, bacteria, cytokines and other reactants are represented as individual 'agents', which behave independently according to a set of predetermined rules which reflect their behaviour in the biological system. The rules for each agent define how they move in space (either in continuous space or on a lattice grid), how they interact with other agents, and when they might be introduced to or removed from the system altogether. In general, these rules are used to update the location and status of each agent at discrete time steps. In some cases, 'hybrid' models are used, in which some reactants are represented as agents, and others are described as background environmental factors controlled by ODEs or PDEs, which may affect the behaviour of agents. This is particularly common for substrates such as chemokines, and the extracellular matrix (ECM) if this is included in the model.

Similarly to discrete stochastic models, ABMs may have the advantage of accuracy at small reactant counts over continuous, deterministic ODE and PDE approaches; ABMs may also be more suitable for representing behaviour such as fluctuations and switching, if this arises.

Another major advantage of ABMs is their flexibilty; they can easily be adapted to include different behaviours for agents simply by updating the defining rules. It is also possible to include far more complex behaviours than are practical to represent in an ODE or PDE; for example, explicitly including the biomechanics of cells (this may include changes of dynamics due to deformation, collision, adhesion or interaction with the ECM). It is also much more feasible to combine modelling approaches across scales using ABMs; a number of ABMs have been developed which integrate the results of inflammation at an organ level to understand the effects at a whole-organism level [An, 2001, Cockrell and An, 2017, McDaniel et al., 2019]. However, the flexibility of ABMs comes at a cost; they are generally time consuming and computationally expensive to run (although they are usually easily parallelised), especially in comparison to most ODE or PDE solvers. Construction of an ABM generally relies on a large number of assumptions, including parameter estimates, as they often include

processes which have not been quantified based on experimental data. Furthermore, they lack the transparency of ODE and PDE models, and the results are not easily generalised to different parametrisations, making them difficult to interpret.

ABMs have been used in numerous contexts related to inflammation and injury, including vocal fold inflammation [Li et al., 2008, Seekhao et al., 2018] and ulcer formation in spinal cord injury [Solovyev et al., 2013, Ziraldo et al., 2015]. Other models have been created to represent chronic pathologies including fibrosis of the liver [Dutta-Moscato et al., 2014], asthma [Song et al., 2012, Pothen et al., 2015], pulmonary fibrosis [Warsinske et al., 2016] and chronic obstructive pulmonary disorder [Ceresa et al., 2018]. Other models have linked the effects of inflammation with oncogenesis [An and Kulkarni, 2015]. As with ODE and PDE models, however, most models describe generic inflammation or specific bacterial infections; aside from *Salmonella* infection, models have been created specifically to describe the effects of *Mycobacterium tuberculosis* [Fallahi-Sichani et al., 2010, Cilfone et al., 2013, Marino and Kirschner, 2016], *Heliobacter pyri* [Carbo et al., 2013, Alam et al., 2015], and *Pseudomonas aeruginosa* [Seal et al., 2011].

One ABM particularly relevant to lesion formation was developed by Shi *et al.* [Shi et al., 2016c, Shi et al., 2016b]. The aim of this study was to analyse how innate and adaptive immune processes link to organ dysfunction, persistent infection or healing in *Salmonella* infection in the liver, and to provide a tool to trial potential therapeutic interventions. This ABM includes many classes of agents including hepatocytes, Kupffer cells, bacteria, resting and active neutrophils, TNF-α, HMGB-1, monocytes, monocyte-derived macrophages and IL-10. This model captured much of the complexity of inflammation resulting from *Salmonella* infection, and unusually, a complex spatial domain was constructed to represent the structure of liver tissue. However, although many parameter values were derived from existing experimental literature, in many cases approximations or estimations had to be used to fully parametrise the model. Furthermore, instead of analysing the spatial behaviour of the system, much of the analysis in this model focused on derived global variables such as total counts for different classes of agents across the domain. This therefore calls into question how necessary the construction of a biologically accurate spatial structure was in this analysis. It is also worth noting that this model does not include phagocytosis of bacteria or lysis of infected cells.

More recently, a hybrid ABM of general inflammation on a tissue level was proposed

by Bayani *et al.* [Bayani et al., 2020b]. In this model, macrophages and neutrophils are modelled as agents, while background levels of pro- and anti-inflammatory mediators are modelled using spatial PDEs. Significantly, results from this model were linked successfully to *in vitro* cell tracking data, demonstrating that the ABM succesfully captured the movement of individual cells; the model was also used to demonstrate how aberrant chemotaxis in neutrophils can lead to failure of resolution of inflammation. However, once again there was little consideration of spatial pattern formation in this study, with a focus on overall abundance of inflammatory markers over time.

Overall, ABMs capture the complexity of biological systems extremely well, and are particularly useful tools both in a clinical or *in silico* experimental context, and are much more likely to have predictive power than the ODE and PDE methods described above. However, they do not necessarily give more insight into the aspects of a system which are particularly responsible for its characteristic behaviour. While conclusions can still be drawn by statistical analysis of simulation output (which is in fact similar to the approach used in Chapter 5 of this thesis, albeit in a different context), it is worth considering whether, for some purposes, a less computationally-intensive method such as spatial PDE modelling might in fact give more insight.

It is also worth noting that, once again, none of the models described here give a clear representation of bacterial infection and intracellular reproduction dynamics, nor is lysis explicitly included in the models described above.

1.8 Research aims

As outlined above, there are many unanswered questions and inconsistencies surrounding both inflammasome formation processes and lesion formation. The overall aim of this research with regard to the inflammasome is to provide a more cohesive answer to the question of how the inflammasome forms. In particular, I wished to address the following points:

- Where does variation in inflammasome formation between samples arise?
- Why is there variation between the behaviours of different inflammasome systems?
- What form do NLR oligomers take?
- How many NLR oligomers are involved in the formation of a single inflammasome?

- Do ASC specks form through linear filamentous growth of PYD-PYD filaments, or does branching of these filaments occur?

Answering these questions using a coagulation modelling methodology will allow us to build towards a logically cohesive conceptual model of inflammasome formation.

The aim of this research with regard to lesion formation is to clarify why the distinct forms of hepatic lesions resulting from *Salmonella* infection arise; most importantly, why lesion size is stable over time, why there is very little mixing of cells and bacteria between lesions, and why lesions are limited in size even as the bacterial population grows.

In order to do this, I will use a simple spatial PDE model of inflammation. I have chosen to use this approach since it permits analysis of spatial structure without the complexity of an ABM. In addition, since insight into the system rather than predictive power is useful here, an ABM would be less appropriate. Unlike previous studies, the focus here will be on spatial structure and pattern formation, rather than the total number of reactants at late times and characterisation of late-time behaviour corresponding to different clinical outcomes.

Furthermore, since it is rare for models to include the downstream effects of inflammasome formation, including lytic cell death and reduction of motility in macrophages following infection, these will all be built into the model. This will allow us to examine how these processes affect lesion formation in more detail. In particular, I aim to establish how the balance of apoptosis and lysis, and the decrease in motility of phagocytes following infection, contribute to lesion formation.

Chapter 2

Analysis of an experimental dataset describing inflammasome formation

2.1 Introduction

Any successful model of the inflammasome must be created in the context of the existing experimental data. While this data is abundant, much is either qualitative, or fairly sparse in terms of data points, rendering it inappropriate for statistical analysis or model fitting. Some of the most mathematically useful datasets involve time series microscopy of cells with fluorescently labelled proteins, similarly to the dataset described in [Cheng et al., 2010], for transfected ASC in HeLa cells.

In this chapter I will present two existing experimental datasets of inflammasome formation, which are based on time series imaging of cells forming NLRP3 inflammasomes. I will demonstrate how quantitative data describing abundance of reactants involved in inflammasome formation can be extracted from these images, as well as explaining the shortcomings and difficulties of analysing this dataset. I will then describe various methods of extracting summary statistics from the quantitative data, which will be analysed in its own right, and later compared to mathematical models (see Chapters 4 and 5). I will also discuss how these datasets motivate model development. Finally, I will explain possible future extensions to the analysis of these datasets, both mathematical and experimental.

2.2 Description of experimental dataset

The two datasets I have used are similar to the datasets presented in the previous study carried out by Cheng *et al.* [Cheng et al., 2010], but represent ASC speck formation as a result of NLRP3 inflammasome activation. These datasets can be used to derive relative abundances of ASC monomers over time, as well as ascertaining the time of ASC speck formation, in cells showing inflammasome formation.

In order to be able to carry out statistical analyses on these datasets, I summarise the relevant time series data for each cell by deriving a number of 'characteristic times', which describe biologically meaningful properties of each time series; for example, time at which an ASC speck forms, and time taken for monomer abundance to be reduced by a given fraction in both compartments; these are explained more fully in Sections 2.3.3 and 2.3.4. This 'characteristic time' approach allows easy comparison of time series within the dataset, so that conclusions may be drawn about variation in inflammasome formation within a population. Analogous characteristic times can also be derived from the solutions to mathematical models, giving us a direct point of comparison between experimental data and models.

Both experimental datasets used were generated by Alessandro Rizzo, and consist of live-imaging time series of murine macrophages from a cell line (C57/BL6 NLRP3-/- NLRP3-flag ASC-mCerulean) which overexpresses ASC with a cerulean fluorescent tag [Andrea et al., 2013]. In one sample, caspase-1 has also been labelled with FLICA, and in the other, MitoTracker has been used to image the mitochondria of the cell. However, I primarily focused on the behaviour of ASC, which is isolated in the cerulean channel (the varieties of MitoTracker and FLICA used both appear in the red channel).

The cells were seeded at a density of 2×10^8 cell/ml in a glass-bottomed dish and the following day were primed with 200 ng/ml LPS for 3 hours. In the case of the MitoTracker data, at this point the cells were stained with MitoTracker Red CMXRos. The cells were then treated with 10 μl nigericin immediately prior to imaging (in the case of the FLICA dataset, 1.5 μl of FLICA 660-YVAD-FMK was dissolved in the media). This stimulated formation of the NLRP3 inflammasome in the cells in the sample.

For the FLICA dataset, imaging was carried out on a Leica SP5 multi-channel AOBS confocal laser scanning microscope, equipped with a temperature and CO2-controlled sample chamber for live-cell imaging, zoom 2x, laser 458 50%, 633 20%, objective HCX PL

APO CS 40.0x1.25 OIL UV, with images taken at intervals of 5.16 seconds. For the MitoTracker dataset, imaging was carried out on an Olympus FV1200 microscope equipped with a temperature and CO2-controlled sample chamber for live-cell imaging, zoom 2x, laser 458 1.5%, 559 0.5%, objective PLAPON60XOSC2 NA:1.40 with images taken at intervals of 1.69 seconds.

The FLICA dataset consists of four time series taken over three independent experiments, and the MitoTracker dataset consists of eleven time series taken over five independent experiments, with a variety of different endpoints.

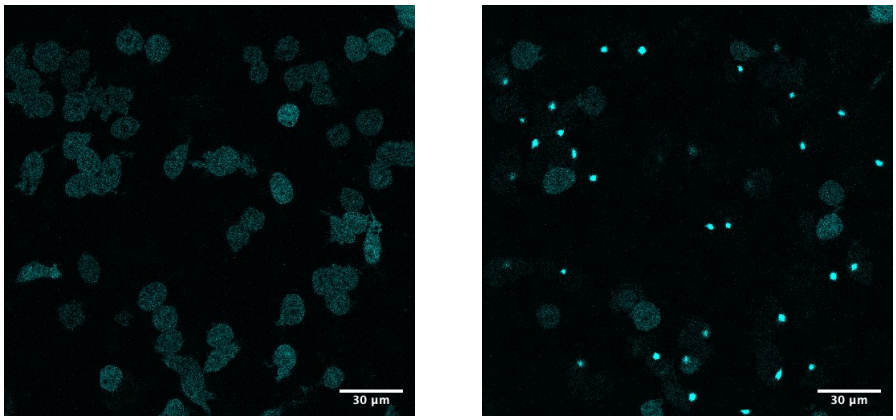
In the resulting images, fluorescence in the cerulean channel is initially diffuse throughout the cell, both in the nucleus and cytosol (Figures 2.1 and 2.2). Following activation, the total intensity in the nucleus and cytosol decreases almost to background levels, with depletion of intensity faster in the cytosol than the nucleus, while intensity in the speck area sharply increases as the inflammasome was assembled.

For each cell in these images, I therefore tracked the total intensity of the cerulean channel in the nucleus, cytosol, and the area where the ASC speck finally formed, over time. These intensity measurements are arbitrary and cannot be used to derive the exact concentrations or abundances of ASC in the cell; however, they can be used as a proxy for the total abundance of ASC monomers in each region relative to other regions, and the initial intensities. I therefore extracted quantitative data from the FLICA and MitoTracker datasets using a custom-designed, semi-automated MATLAB script, loosely based on code written by Alessandro Rizzo and Eugenia Cammarota. I will describe this analysis in the following section.

2.3 Analysis of data

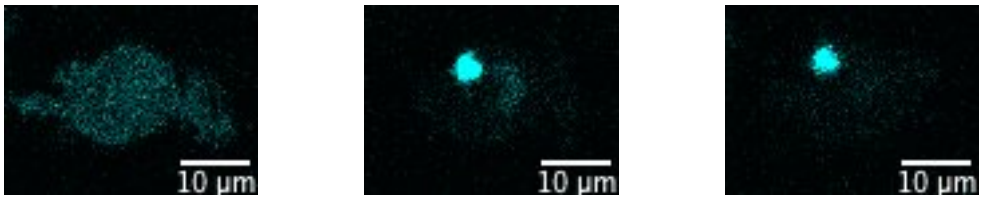
2.3.1 Preprocessing

Since there are many frame-wide discontinuities in intensity across time series, I normalised intensities in each frame by dividing by the total intensity across the frame at each timepoint. This had a smoothing effect on output data, which was then more amenable to fitting.



$t = 0s$ $t = 1700s$

Figure 2.1: First and last frame of image from the FLICA dataset (Well 1, 14/05/2014) shown in the cerulean channel; time is measured from the addition of LPS. ASC appears as cerulean in these images; initially ASC is diffuse throughout the cells. In the final frame, some cells display a very bright, concentrated spot corresponding to the ASC speck, indicating that inflammasome formation has taken place; others still show ASC diffuse throughout the cell, indicating that inflammasome formation has not taken place.



$t = 0s$ $t = 135s$ $t = 290s$

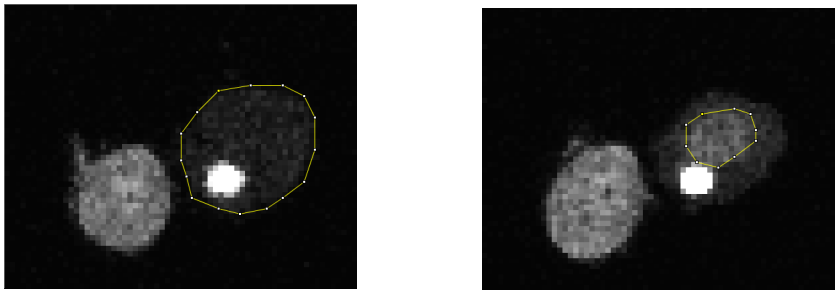
Figure 2.2: Example frames cropped to show one cell from the dataset in Figure 2.1. ASC is initially diffuse throughout the cell. By $t = 135s$ after infection, an ASC speck is visible, but not all ASC has been depleted from the rest of the cell; the nucleus, from which ASC is depleted more slowly, is visible as a shadow adjacent to the speck. By $t = 290s$, virtually all ASC throughout the cell has been depleted.

2.3.2 Extraction of intensity traces

In order to isolate the cell and nucleus areas, the script requires submission of a user-defined region of interest (ROI) using a point-and-click method to delineate a polygon on a user-specified frame in which the relevant area can be seen clearly (Figure 2.3). While this method requires considerable input from an individual user, it is generally robust to variations in frame choice and ROI definition. I analysed only cells which remain fairly static, remain fully within the field of view throughout the time series, do not overlap with other cells, and display a fully-formed speck before the time series concludes. In total, I identified 53 usable cells in the FLICA dataset and 23 usable cells in the MitoTracker dataset All cytosolic and nuclear traces are shown in Figure 2.4.

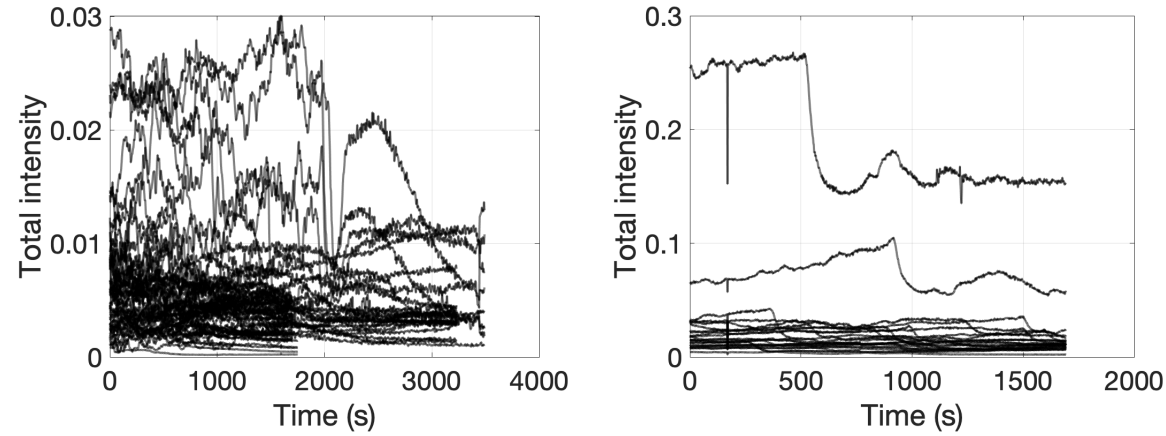
Following user submission of cell and nucleus areas, the speck area is defined automatically by thresholding, for a user-specified frame in which it is clearly visible; setting a threshold of 85% of the maximum intensity within the cell area in the given frame was sufficient to isolate the speck in all cases. The intensities of pixels within each of these ROIs (cytosol, nucleus and speck) are then integrated over time to give individual intensity traces (Figure 2.5).

In general, the traces for cytosolic and nuclear regions show an initial lag in which the intensity remains approximately constant, followed by a period of sharp decrease to a plateau intensity at background level, as ASC monomers are absorbed into the ASC speck. The initial lag is likely to originate from various processes, including time taken for nigericin to be internalised by the cells, and for activation (and possibly oligomerisation) of NLRP3 to take place. There is considerable variation in the length of this lag, even within the same sample. There is also variation in the gradient of the intensity decrease in nuclear and cytosolic traces corresponding to monomer depletion by speck formation. The speck region shows an initial lag followed by a sharp increase as the ASC speck forms, finally reaching a plateau as the pixels become saturated. All traces show considerable noise due to random fluctuations in intensity. Not all cells within the dataset show speck formation before the endpoint of the timeseries; however, it is useful to track the number of cells which do not form specks during the time series, to gain a fuller view of the distribution of speck formation times.

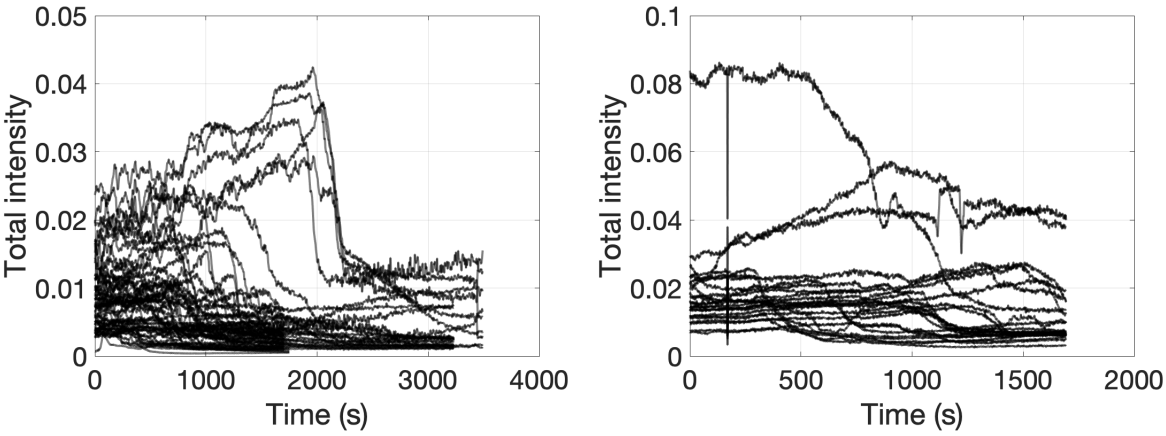


Cell ROI Nucleus ROI

Figure 2.3: Examples of ROI delineation to demarcate the cell and nucleus areas (shown as yellow polygons) for a cell from the FLICA dataset.

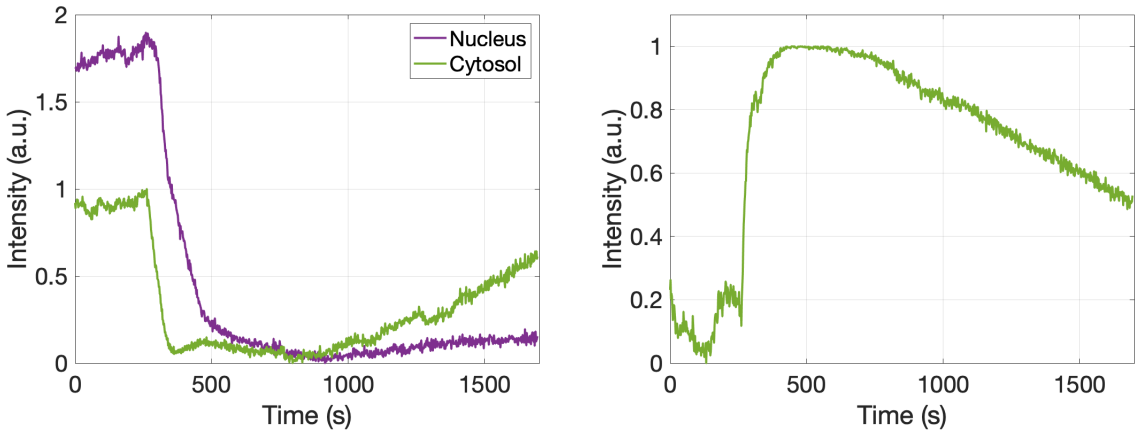


Cytosolic traces, FLICA dataset Cytosolic traces, MitoTracker dataset



Nucleus traces, FLICA dataset Nucleus traces, MitoTracker dataset

Figure 2.4: Plots of all traces from the FLICA and MitoTracker datasets in the nuclear and cytosolic regions



Cytosolic and nuclear traces

Speck area trace

Figure 2.5: Representative examples of traces extracted from data for one cell from the MitoTracker dataset. Note that all traces have been scaled by the total intensity in each frame. Following an initial plateau, the cytosolic and nuclear traces show a decrease corresponding to ASC monomer depletion, followed by another plateau as no ASC monomers remain. After the initial plateau, the speck trace shows a rapid increase as the ASC speck grows, which reaches another plateau as the pixels in this area become saturated.

2.3.3 Fitting breakpoints

In order to understand when inflammasome formation begins and monomers begin to be depleted, I aimed to divide the traces into two periods over which the intensity is approximately constant (the initial lag and final plateau), separated by a period of intensity decrease for the cytosolic and nuclear regions, and a period of increase in the speck region. Identifying the onset of intensity decrease in the speck and nuclear regions gives an estimate for the time at which inflammasome formation begins; meanwhile, identifying the time at which the speck region reaches a plateau gives an estimate for the time of completion of inflammasome formation. Identifying the start and end of the decrease in intensity in the nuclear and cytosolic regions also gives an interval for which a rate of decrease can be fitted.

In order to identify the 'breakpoints' between these regions, I fitted a trilinear model to each trace, i.e., a continuous curve consisting of three straight lines with different gradients; an example is shown in Figure 2.6. I assumed that the gradient in each trace before and after inflammasome formation is approximately constant, but need not be zero (this may help to account for the photobleaching effect). Similar methods have been used elsewhere to identify changepoints in the gradients of noisy data, for example, in identifying epidemic onset from disease prevalence data [Charu et al., 2017]. It is worth bearing in mind that the initial lag may partly arise through other processes (for example, NLR oligomerization)

taking place before cluster joining and growth begins; however, the models demonstrate that the depletion of monomers and growth of clusters may follow sigmoid kinetics, which may also contribute to the initial lag. This will be examined further in Chapter 6.

I have assumed that the ith datapoint x_i is approximately distributed as follows:

$$x_i = \beta_0 + \beta_1 t_i + \beta_2 (t_i - t_{b_1})^+ + \beta_3 (t_i - t_{b_2})^+ + \epsilon_i \tag{2.1}$$

where t_i is the time at which measurement x_i is taken, t_{b_1} and t_{b_2} are the start and end point of the interval of speck formation, β_0 is the initial intensity, and β_1, β_2 and β_3 are the gradients of each section of the trilinear fit. ϵ_i is the error at datapoint i; I have assumed these are independent and normally distributed with the same variance, which allows the model to be treated as a standard linear model.

$(t_i - t_{b_j})^+$ is defined as follows:

$$(t_i - t_{b_j})^+ = \begin{cases} t_i - t_{b_j} & \text{if } t_i > t_{b_j} \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

I fitted this model to all traces, calculating the maximum likelihood for each possible pair of t_{b_1} and t_{b_2} by fitting a standard linear model using the least squares method. I then estimated the values of t_{b_1} and t_{b_2} by choosing the pair yielding the largest overall maximum likelihood.

In some cases, I excluded cells from further analysis since the code failed to fit one or more traces associated with that cell; this generally took place when speck formation occurred within a few frames either of the start or finish of the time series, so that either the first or last section of the trilinear fit was extremely short, or the noise around these points prevented identification of the breakpoint in question by the algorithm. I also excluded traces for the nuclear and cytosolic regions which did not show a negative gradient between the two breakpoints.

For both datasets, the trilinear fit model successfully identified intervals of speck formation with negative gradient for cytosolic and nuclear traces for only 34/53 FLICA traces and 19/23 MitoTracker traces; however, it was more consistently successful than any other method trialled (for example, using separate bilinear fits to identify both breakpoints). The trilinear fit method is also preferable since it does not require constraints on the β values, considers the whole trace rather than just a small interval, and requires no user input.

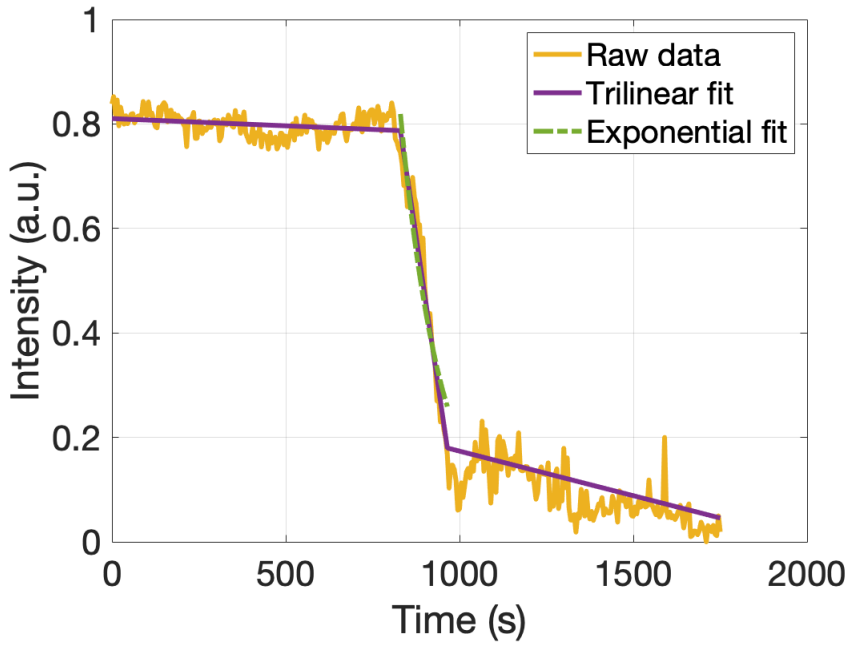


Figure 2.6: Trace extracted for the nuclear compartment for one cell from the FLICA dataset, with trilinear fit and fitted exponential decrease for that trace plotted in purple and green respectively.

Once I had identified the interval of speck formation in nuclear and cytosolic traces, I then fitted a single-term exponential function of the form ae^{bt} to this interval, to quantify the rate of decrease in the cytosolic and nuclear traces during speck formation; an example is given in Figure 2.3.4.

2.3.4 Estimation of monomer depletion characteristic times

Another set of variables which can easily be compared to simulated data and models are the monomer depletion characteristic times $t_{1/q}$ — the time taken for each monomer population to deplete by a fraction $1/q$ of its initial value. For each usable trace, I therefore estimated the value of $t_{1/2}$ and $t_{1/10}$. To avoid noise due to random fluctuations in intensity, I first smoothed the data by calculating a moving median for each point in the time series, using a window size of 10 (this was sufficient to smooth noise in the original dataset without losing too much resolution). I then located the first timepoint after t_{b_1} when the intensity was less than $(1-1/q)\beta_0$, for the values of t_{b_1} and β_0 calculated earlier. An example is given in Figure 2.7. I also excluded traces for which $t_{1/2}$ and $t_{1/10}$ could not be identified, leaving 30 traces in the FLICA dataset and 20 in the Mitotracker dataset.

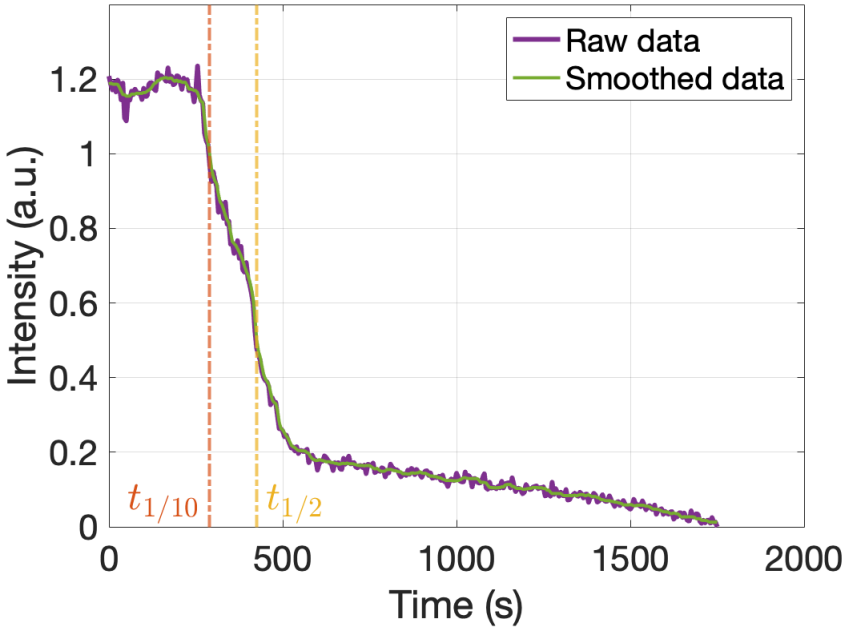


Figure 2.7: Example of total intensity in the nuclear region for a cell in the FLICA dataset, scaled by total intensity across the frame (plotted in purple), with smoothed data plotted in green above. Characteristic times for this trace are also shown: $t_{1/q}$ represents the time taken for the intensity to decrease by a fraction $1/q$.

2.3.5 Limitations

Even post-normalisation, the trace extraction method described above inevitably results in somewhat noisy data; while cells are theoretically fairly static throughout the time course, especially in cells in which inflammasomes have already formed (see [Man et al., 2014a]), there is still some movement on the whole-cell level. Cells may move laterally, out of the user-designated ROI, and in some cases, enter the ROI of other cells. In some rare cases, the cells may move out of the focal plane, so that the speck is no longer visible. Further inaccuracies are caused by the phenomenon of photobleaching, which results in a decrease of intensity of individual fluorophores over time; this is particularly evident in many traces which demonstrate an initial slight decrease in intensity well before speck formation or in the plateau following speck formation (Figure 2.5). Furthermore, the information which can be extracted from the trace for the speck area is limited due to the fact that the pixels in this area tend to saturate before cytosolic and nuclear intensities have fully depleted.

It is also important to note that macrophages are three-dimensional entities; I have assumed that the two-dimensional images in this dataset form a representative slice through the cell, and that there is negligible variation in abundances in the z-direction, but this cannot

be guaranteed. A further source of inaccuracy arises from the fact that the speck area is removed from the cytosolic compartment before the speck itself is formed; however, this is sufficiently small that it is unlikely to make a significant difference to our overall results.

Finally, it is worth bearing in mind that the data is truncated by the endpoint of the time series of images; ASC speck formation times cannot be identified past this time, although in practice the truncation time may be somewhat earlier, since breakpoints cannot be identified if speck formation happens very close to the end of the time series. Other characteristic times cannot be identified for cells with sufficiently late ASC speck formation times. It is impossible to distinguish, on the basis of this dataset, whether the cells that do not form specks will never do so, or would do so at timepoints beyond the end of the time series; however, cells do not have an indefinite window for inflammasome formation to take place, since other biological processes will eventually intervene. Therefore, cells which would theoretically fall in the upper end of the distribution of speck formation times will not in fact ever demonstrate speck formation.

A summary of endpoint times and the number of cells analysed for each time series is given in Table 2.1.

2.4 Results

2.4.1 Estimation of relative protein abundances in nucleus/cytosolic region

I also used nuclear and cytosolic traces to estimate the relative abundance of monomers in the nuclear and cytosolic regions at equilibrium; this is a useful guideline for simulations and modelling. I took the ratio of abundances to be equal to the ratios of the respective fitted β_0s for nuclear and cytosolic traces. Since there appeared to be a number of outliers in both datasets, I took $\frac{\hat{x}_n(0)}{\hat{x}_c(0)}$ to be the median observed value of the ratio of β_0 for the nuclear and cytosolic traces. This gave $\frac{\hat{x}_n(0)}{\hat{x}_c(0)} = 1.48$ for the FLICA dataset, and $\frac{\hat{x}_n(0)}{\hat{x}_c(0)} = 1.11$ for the MitoTracker dataset.

2.4.2 Comparison of datasets

Using a Mann-Whitney U test shows that there are no significant differences between the two datasets for any of the following statistics: t_{b_1}, t_{b_2}, decay constants, $t_{1/2}$ and $t_{1/10}$ for cytosolic and nuclear traces and t_{b_1}, t_{b_2} for speck traces; p-values are given in Table 2.2.

Dataset	Date	Sample	Max time (s)	Cells without speck	Cells with speck	Traces analysed
FLICA	12/05/2014	Well 2	1750	42	12	4
FLICA	14/05/2014	Well 1	1700	19	40	9
FLICA	14/05/2014	Well 2	3230	5	29	6
FLICA	26/05/2014	Well 1	3490	10	14	5
MitoTracker	07/08/2015	Image0006	NA	NA	NA	NA
MitoTracker	11/08/2015	Image0008	2500	4	1	0
MitoTracker	11/08/2015	well_2	5000	6	5	3
MitoTracker	12/08/2015	Aug-15	5000	5	3	0
MitoTracker	12/08/2015	Image0005	NA	NA	NA	NA
MitoTracker	13/08/2015	Image01_01	NA	NA	NA	NA
MitoTracker	13/08/2015	Image0002	NA	NA	NA	NA
MitoTracker	13/08/2015	Image0005	5000	2	2	0
MitoTracker	13/08/2015	Image0007	5000	1	1	1
MitoTracker	14/08/2015	Image0003	5000	2	18	13
MitoTracker	14/08/2015	Image0005	2335	3	2	0

Table 2.1: Traces analysed from the FLICA and MitoTracker datasets. Samples with no speck-forming cells appropriate for analysis have been excluded. The number of cells showing a speck by the end of the time series, as well as the number of cells which do not show a speck, are given for each sample image. The number of traces which were successfully processed for each image are also given (i.e. traces in which cytosolic, nuclear and speck region traces were all successfully fitted, with a negative gradient in the second of the three continuous straight line segments forming the trilinear fit).

Statistic	p-value
Decay constant (cytosol)	0.5516
Breaktime 1 (cytosol)	0.7209
Breaktime 2 (cytosol)	0.9894
$t_{1/2}$ (cytosol)	0.4668
$t_{1/10}$ (cytosol)	0.7609
Decay constant (nucleus)	0.5516
Breaktime 1 (nucleus)	0.3613
Breaktime 2 (nucleus)	0.8634
$t_{1/2}$ (nucleus)	0.6433
$t_{1/10}$ (nucleus)	0.6623
Breaktime 1 (speck)	0.2607
Breaktime 2 (speck)	0.4349

Table 2.2: p-values when comparing statistics derived from FLICA and MitoTracker datasets respectively using a two-tailed Mann-Whitney U test to test the null hypothesis that the statistics have the same distribution for the two datasets. There is no significant difference between the datasets at the 0.05 level.

Therefore, I have pooled the results from the two datasets; although the methods of data acquisition were slightly different and different molecular tags in the preparation were present in the cells, this does not appear to have had a material effect on results for the cerulean channel. Furthermore, despite the variation in time series endpoints, many of the speck formation times are concentrated at earlier timepoints, so the distributions are affected very little.

2.4.3 Comparison of results in nuclear and cytosolic regions

Using a paired one-sided Wilcoxon signed rank test to compare decrease onset times for nuclear and cytosolic intensity traces shows that nuclear traces have later decrease onset times (t_{b_1}) with $p = 2.6367 \times 10^{-4}$ for the pooled datasets. Likewise, exponential decay constants are greater for cytosolic data with $p = 5.9390 \times 10^{-5}$ (Figures 2.8 and 2.9). Unsurprisingly, $t_{1/2}$ and $t_{1/10}$ are significantly larger for nuclear traces compared to cytosolic traces ($p = 2.0152 \times 10^{-7}$ and $p = 2.8357 \times 10^{-7}$respectively). This suggests that in general, depletion of ASC monomer from the nuclear compartment occurs later and is slower than depletion of monomers from the cytosolic compartment. This supports the assumption that the inflammasome forms in the cytosolic compartment, and that no clusters are present within the nucleus, since monomers in the nuclear compartment must first cross into the cytosolic compartment before they can be absorbed by a cluster.

Notably, the distributions of all breaktimes and both $t_{1/2}$ and $t_{1/10}$ have a strong positive skewness; this suggests that most cells that form inflammasomes within the time series do so at early timepoints, but also that there is a spread of inflammasome formation occurring at much later timepoints for a minority of cells.

2.5 Discussion

2.5.1 Summary of results

In this chapter I have described two typical imaging datasets showing the formation of NLRP3 inflammasomes, and presented a method to extract and then analyse quantitative data from the image time series. I have demonstrated how summary statistics can be obtained from the resultant quantitative datasets, or 'traces'. These include 'breakpoints' estimating when

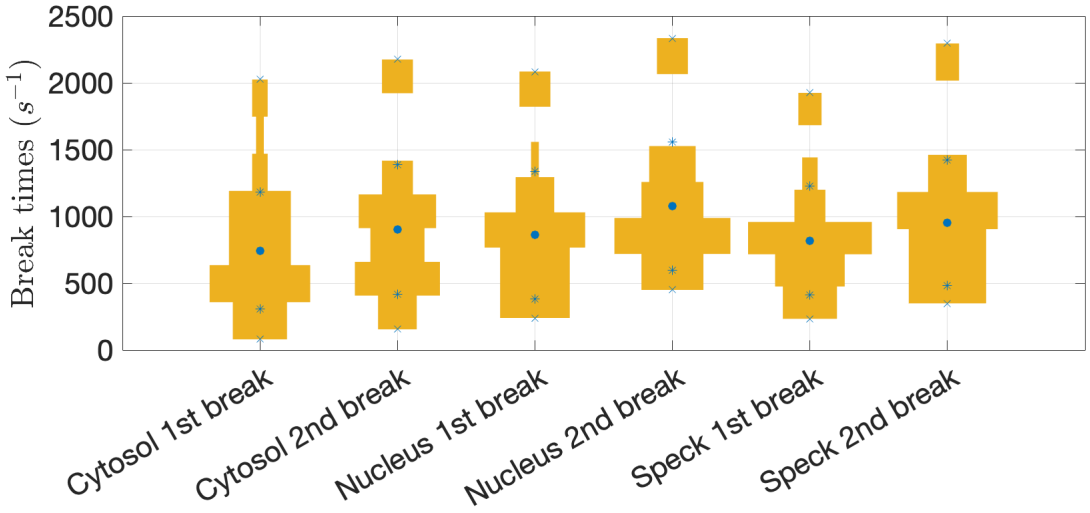


Figure 2.8: First and second breakpoint times for nuclear and cytosolic traces fitted to trilinear model with pooled datasets. Mean is shown with a circle, mean $\pm$ standard deviation is marked with a star, and maximum and minimum values are marked with a cross. Depletion from nuclear regions occurs later than in cytosolic regions.

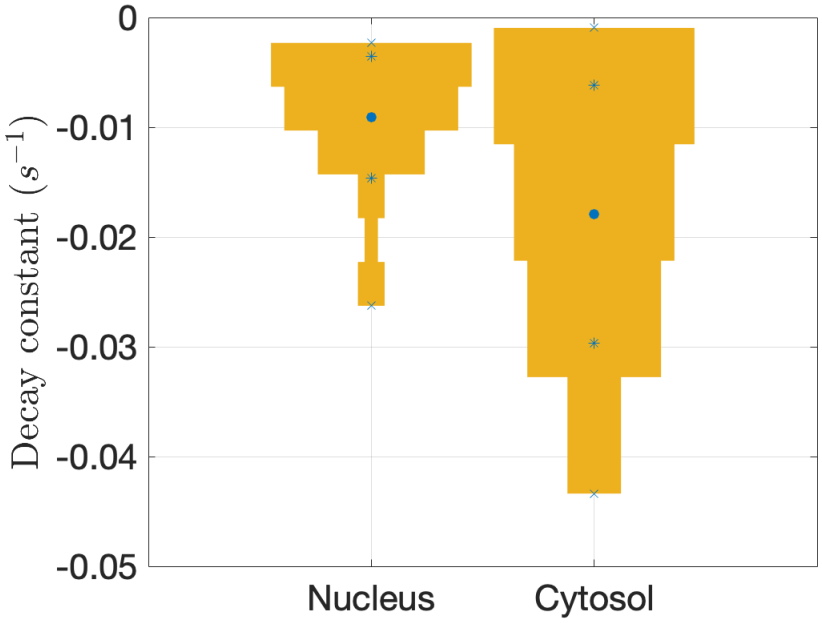
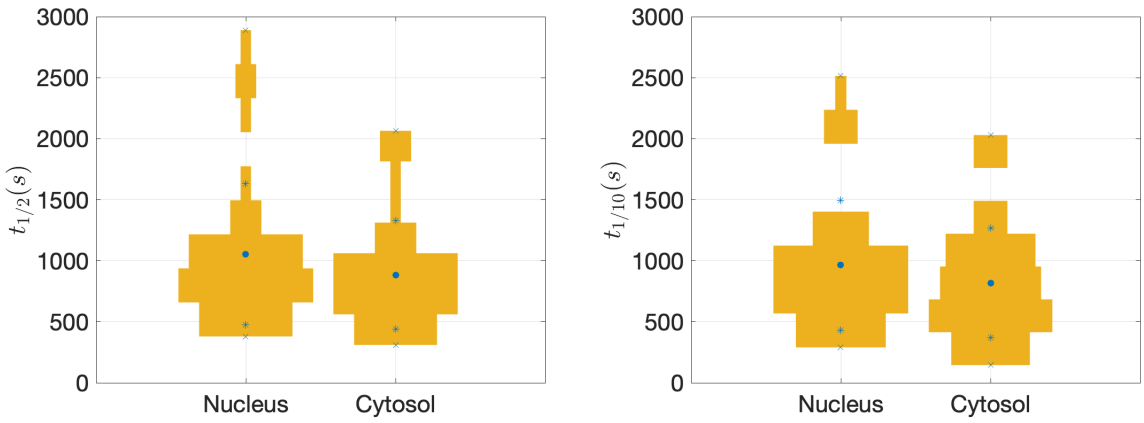


Figure 2.9: Exponential decay constants, for nuclear and cytosolic traces fitted to trilinear model with pooled datasets. Mean is shown with a circle, mean $\pm$ standard deviation is marked with a star, and maximum and minimum values are marked with a cross. Depletion from nuclear regions is slower than from cytosolic regions.



(a) $t_{1/2}$ (b) $t_{1/10}$

Figure 2.10: Monomer depletion characteristic times for nuclear and cytosolic traces for pooled datasets. Mean is shown with a circle, mean $\pm$ standard deviation is marked with a star, and maximum and minimum values are marked with a cross. Depletion from nuclear regions is slower than from cytosolic regions.

depletion of ASC monomers from nuclear and cytosolic regions begins and ends and analogous onset times for the emergence of the ASC speck; estimates for the rate of decrease in monomer abundance; and the monomer depletion characteristic times $t_{1/q}$, which describe the time taken for the ASC monomer abundance in a region to drop by a fraction $1/q$.

These statistics are useful for comparison with models and other datasets, but also valuable in their own right. In particular, I have demonstrated that depletion of ASC monomers from the nuclear region of the cell occurs more slowly and later than depletion from the cytosolic region.

2.5.2 Motivation for model development

Since this dataset tracks the relative abundance of ASC monomers over time, this naturally suggests a model which similarly tracks the abundance of ASC monomers and NLRs over time, similar to the aggregation models described in Chapter 1. While the data only captures the dynamics of ASC monomers, and to some extent the emergence of the ASC speck, the mathematical model will certainly have to include other variables not included in the dataset (for example, abundance of NLR oligomers, and clusters of various sizes).

The dataset also suggests key attributes which we would expect a successful model to display; in particular, we expect to see ASC monomer abundances gradually declining to zero, alongside rapid emergence of a single ASC speck. As demonstrated, the dynamics

of monomer depletion from the nuclear and cytosolic regions of the cell are clearly different, and even if the model is not fully spatial, it would be useful to include a division of the ASC monomer population between these two compartments.

The model must be constructed with these essential properties in mind. It will also be important to determine the values of various summary statistics from the mathematical model which can be compared to those derived from the experimental dataset in this chapter (for example the monomer depletion characteristic times $t_{1/2}$ and $t_{1/10}$, and the time taken for ASC specks to form); this will allow a more quantitative comparison of the model to the experimental data.

The resulting models will be discussed in the following chapter.

2.5.3 Future work

The analysis in this chapter has a number of limitations, which could be addressed in future mathematical and experimental work. There were some key limitations with the dataset, perhaps most importantly the limited number of usable traces, and the limited number of both technical and biological replicates. This reduces the robustness of our results, makes it much harder to establish clear trends or fit distributions to the data, and will also pose difficulties when comparing to mathematical models (this will be examined further in Chapter 6). To add to the difficulty in analysing these datasets, there was also considerable inconsistency in the lengths of the time series and the frame rates used. The large degree of noise in each of the traces also led to difficulties in fitting many traces, reducing the size of the usable dataset even further. Therefore, a key priority in continuing the work of inflammasome modelling must be the creation of more, larger, and more consistent datasets. The analysis described in this chapter could then be easily repeated.

A possible extension would be creating similar datasets in three dimensions; this would be much more technically challenging than the two dimensional time series imaging described above, but it would help to address the issue of cells and ASC specks moving out of the plane of focus, and also allow us to establish whether our assumption that a two-dimension slice of a cell is representative of all z-values is justified.

Another limitation of working with these existing datasets is that the intensity of fluorophores only gives us a proxy for the abundance of ASC monomers in each region; since concentrations or absolute numbers of ASC monomers are not currently available, this will

further add to the difficulties in modelling inflammasome formation. However, hopefully the creation of a model of inflammasome formation will motivate the need for such measurements to be taken, as well as providing a mathematical framework from which certain parameters such as reaction rates can be derived.

Finally, the existing datasets are limited in that they show only the behaviour of ASC in inflammasome formation (due to the inclusion of FLICA and MitoTracker in the two datasets examined, it would technically be possible to examine the behaviour of Caspase-1 and the mitochondria in relation to inflammasome formation, although the datasets for each of these would be even smaller). In future, it would be useful to include fluorescent tagging of other proteins involved in inflammasome formation, especially the relevant NLR; this would allow us to investigate how NLR activation and oligomerisation affects initial lag times, as well as how they influence aggregation over the course of inflammasome formation. These experiments would similarly be more technically challenging than those carried out to generate the datasets used in this analysis, since it would require multi-channel fluorescence microscopy, and fluorophores would have to be chosen carefully to avoid bleedthrough between channels. Furthermore, it may be necessary to use different cell lines, and the results may not be directly comparable. However, this would give us an extremely useful insight into the interaction between inflammasome constitutents, which we must currently infer from ASC dynamics only. It would also be useful to conduct similar experiments with other varieties of inflammasome (the NLRC4 inflammasome would be a particularly interesting contrast) so that the results could be compared; this would give a useful insight into the differences in behaviour between inflammasomes. The analysis pipeline established in this chapter could easily be adapted to any of these scenarios.

Further extension to the processing pipeline described in this chapter would be useful; there is much potential in this problem for moving beyond the user input-reliant approach described here. In particular, machine learning techniques could be used as a more sophisticated approach to tracking cells between frames; this has the advantages of requiring far less user supervision, as well as removing the issue of cells moving out of the user-established region over the time series. However, this would be somewhat computationally intensive, and require a substantial training data set; these techniques would be much more appropriate once larger datasets are available. It would also be potentially possible to extract detail from these images at a finer spatial resolution than the division of the cell into nuclear

and cytosolic regions; however, this would require finer-resolution images than were available here. If such images were available, it may also be possible to quantify the distribution of sizes of clusters over time, which would allow much more direct comparison to the models which will be presented in the following chapters.

Chapter 3

Development of a model of inflammasome formation

3.1 Introduction

The analysis of the inflammasome formation dataset presented in Chapter 2 motivates the creation of a model of inflammasome formation similar to the models of protein coagulation described in Section 1.5.

In this chapter I will describe two models of inflammasome formation: a 'full' model, which is closer to the underlying biological processes, and a 'simplified' model, which removes some of the more complex aspects of the full model but is significantly more amenable to mathematical analysis. While they have been inspired by models of comparable biological processes (especially prion formation, as described in Section 1.5.2), both models are entirely novel; as far as I have been able to establish, there are no other existing models of inflammasome formation with this level of detail, and these models are also substantially mathematically different to other mathematical models of coagulation processes, due to the combination of multiple spatial compartments, and multiple constituent monomer species with different behaviours.

For both variations, I present a continuous deterministic ODE version of the model, and a discrete stochastic version; these are appropriate in different contexts, and both will be analysed further in later chapters. For both models, I will explain the reactions considered, as well as the simplifications and assumptions included in the model. I will also examine some useful mathematical properties of each model.

Finally, I will outline the 'characteristic times' which will be used to summarise model solutions, and which can be compared to key features of the experimental data derived in the previous chapter.

3.2 Full model of inflammasome formation

3.2.1 Variables and reactions

To explore the process of inflammasome formation, I focus my attention on abundances of the NLR of interest and of ASC, since these proteins drive inflammasome formation. I assume that initially there are a number of NLR oligomers, formed following activation, of given sizes, and a given number of unbound ASC monomers. These monomers and oligomers combine to form clusters. Building on the existing conceptual models of ASC-ASC and ASC-NLR interactions (for example, [Masumoto et al., 2001, Moriya et al., 2005, Lu et al., 2014, Sahillioglu et al., 2014, Vajjhala et al., 2014, Chu et al., 2015, Sanders et al., 2015, Dick et al., 2016, Oroz et al., 2016, Kuri et al., 2017, Li et al., 2018, de Alba, 2019], I assume that each NLR monomer in a cluster can 'seed' an ASC filament, which extends via PYD-PYD interactions.

I include in the model a mechanism which allows clusters to join by crosslinking of these filaments; this is a putative reaction which has not been observed experimentally, and can be included or removed from the model based on one parameter value. Varying this parameter will allow us to establish the effect such a process would have on inflammasome formation, and indeed ascertain whether it is necessary in the system.

I also assume that ASC CARD-CARD reactions occur concurrently with PYD-PYD interactions; since both PYD-PYD and CARD-CARD interactions in isolation can lead to filament formation [de Alba, 2019], this suggests that an individual ASC monomer could form a unit in two separate filaments, one formed via PYD-PYD interactions and one formed via CARD-CARD interactions. This suggests that if CARD-CARD and PYD-PYD interactions do occur concurrently, this could lead to branched structures forming. While there is currently no explicit supporting biochemical evidence for this theory, this would be a natural consequence of concurrent CARD-CARD and PYD-PYD filament-forming reactions, and could explain why branching structures have been observed in images of ASC specks [Kuri et al., 2017]. In order to test this theory,branching reactions can be removed from the model by setting the

corresponding reaction rate to zero.

In order to track the number of possible reactions with each cluster, I track the number of ASC filaments within the cluster, and the number of 'free' NLRs associated with each cluster which have not yet formed the basis of a filament. I also track the total number of ASC monomers in each cluster; this does not play a role in the dynamics of the model, but allows the total size of each cluster as well as just the number of filaments to be tracked.

Neither of these models are explicitly spatial, although I do separate ASC abundances into two separate cellular compartments corresponding to the nucleus and cytosol. ASC is located endogenously in both compartments, but the depletion of ASC in the nucleus during inflammasome formation is generally slower than in the rest of the cell, since the completed inflammasome complex is normally located outside the nucleus, in the perinuclear region [Cheng et al., 2010]. I have also confined NLR oligomers and clusters to the cytosolic compartment.

The variables used in both models are given in Table 3.1, with an illustrative example given in Figure 3.1. The reactions considered are illustrated in Figure 3.2 and formally defined in Table 3.2. The full ODE model is outlined in Section 3.2.4. Propensity functions for reactions in the discrete stochastic model are given in Table 3.2.

I have developed these models for NLRC4 inflammasome formation, although I have not made any assumptions about reaction rates (Reactions C1 and C3 both nominally describe CARD-CARD interactions but I have not assumed c_1 and c_3 are equal - see Table 3.2 for details). This model could therefore be adapted easily to reflect other inflammasome systems; for example, for the NLRP3 inflammasome, similar reactions occur but C1 reactions are PYD-PYD rather than CARD-CARD reactions.

Note the form of the propensity function for reaction C2 when two clusters of the same size react together (Table 3.2); the propensity function changes from $c_2 f i x_{f,n,a} x_{i,j,k}$ for clusters of different sizes to $\frac{1}{2} c_2 f^2 x_{f,n,a}(x_{f,n,a} - 1)$ for clusters of the same size. If we have a reaction involving clusters with sizes (f, n, a) and (i, j, k), where $(f, n, a) \neq (i, j, k)$, then the rate of reaction is proportional to the total possible number of ordered pairings of such clusters, which is $x_{f,n,a} x_{i,j,k}$. However, if the clusters are the same size (f, n, a), both are drawn from the pool of $x_{f,n,a}$ such clusters; there are $\frac{1}{2} x_{f,n,a}(x_{f,n,a} - 1)$ possible pairings, where ordering is not considered since the clusters are functionally identical. This is similar to the correction terms used in BD models (see Equations 1.2 and 1.3).

Variable	Explanation
$\hat{x}_c$	Total abundance of ASC monomers present in cytosol
$\hat{x}_n$	Total abundance of ASC monomers present in nucleus
$x_{f,n,a}$	Total abundance of clusters in cytosolic compartment with f growing ASC filaments, n free NLRC4 monomers, a ASC monomers

Table 3.1: Variables used in ODE and stochastic models. 'Abundance' refers to concentration in ODE models and absolute numbers in stochastic models.

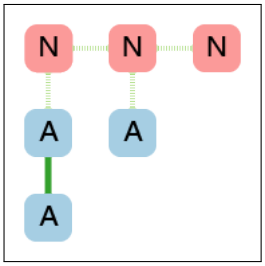
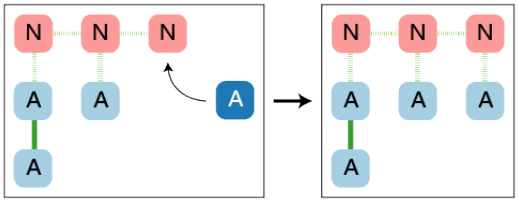


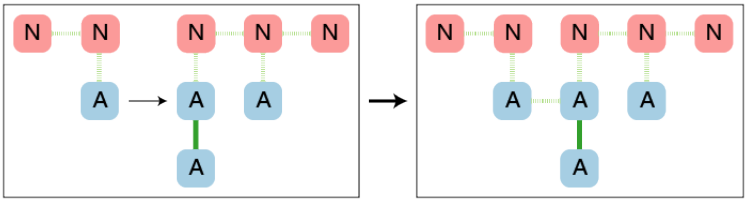
Figure 3.1: The NLRC4 inflammasome in this model: NLRC4 monomers (pink) are joined by CARD-CARD interactions (dashed pale green). Active ASC monomers (pale blue) are either joined to NLRC4 monomers via CARD-CARD interactions, or to other active ASC monomers via PYD-PYD interactions (solid dark green), forming filaments. In the example shown, the cluster has two filaments, one free, unbound NLRC4 monomers, and three ASC monomers; this cluster would count towards $x_{2,1,3}$.

Reference	Reaction	Propensity function
C1	$\hat{x}_c + x_{f,n,a} \xrightarrow{c_1 n} x_{f+1,n-1,a+1}$	$c_1 n \hat{x}_c x_{f,n,a}$
C2	$x_{f,n,a} + x_{i,j,k} \xrightarrow{c_2 fi} x_{f+i,n+j,a+k}$	$c_2 fi x_{f,n,a} x_{i,j,k}$ for $(f,n,a) \neq (i,j,k)$, $\frac{1}{2} c_2 f^2 x_{f,n,a}(x_{f,n,a}-1)$ otherwise
C3	$\hat{x}_c + x_{f,n,a} \xrightarrow{c_3 f} x_{f+1,n,a+1}$	$c_3 f \hat{x}_c x_{f,n,a}$
P	$\hat{x}_c + x_{f,n,a} \xrightarrow{pf} x_{f,n,a+1}$	$pf \hat{x}_c x_{f,n,a}$
M1	$\hat{x}_n \xrightarrow{m_1} \hat{x}_c$	$m_1 \hat{x}_n$
M2	$\hat{x}_c \xrightarrow{m_2} \hat{x}_n$	$m_2 \hat{x}_c$

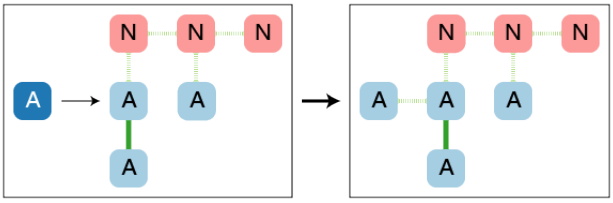
Table 3.2: Rates of reactions in ODE and stochastic models. The reactions are explained more fully in Figure 3.2.



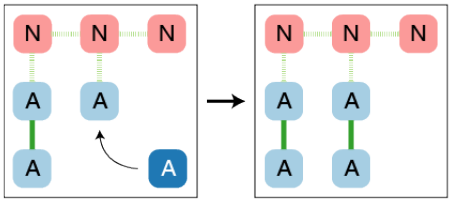
(a) **Reaction C1:** An inactive ASC monomer joins a cluster via a CARD-CARD interaction with NLRC4 in a cluster, and is activated. The number of free NLRC4 in the cluster decreases by one, and the number of filaments increases by one. Here, $x_{2,1,3} + \hat{x}_c \to x_{3,0,4}$.



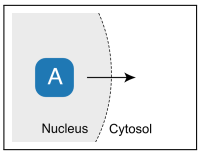
(b) **Reaction C2:** Two clusters join via CARD-CARD interactions between active ASC in a cluster. A new cluster is formed, where the number of free NLRC4 and filaments is the sum of those of the constituent clusters. Here, $x_{1,1,1} + x_{2,1,3} \to x_{3,2,4}$.



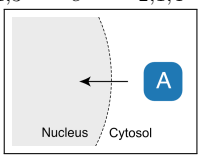
(c) **Reaction C3:** An inactive ASC monomer joins a cluster via a CARD-CARD interaction with active ASC in a cluster, and is activated. The number of filaments in the cluster increases by one. Here, $\hat{x}_c + x_{2,1,3} \to x_{3,1,4}$.



(d) **Reaction P:** An inactive ASC monomer joins a cluster via PYD-PYD interaction with active ASC in a cluster, and is activated. The filament the monomer joins is extended, but the total number of free NLRC4 and filaments in the cluster is unchanged. Here, $x_{2,1,3} + \hat{x}_c \to x_{2,1,4}$



(e) **Reaction M1:** An inactive ASC monomer moves from the nuclear to the cytosolic compartment.



(f) **Reaction M2:** An inactive ASC monomer moves from the cytosolic to the nuclear compartment.

Figure 3.2: Reactions considered in the full inflammasome model

3.2.2 Simplifications and assumptions

It has been necessary to introduce a number of simplifications of, and assumptions about, the biological processes of inflammasome formation when creating this model. Partly this is inherent in the process of model formation, in which we simplify processes which are not relevant to the questions at hand; however, in the case of inflammasome formation, there are also a number of aspects of the biological process which are poorly-understood, or have not been quantified to the extent which would be ideal when creating a model. Therefore, it has been particularly necessary to introduce various assumptions about the process of inflammasome formation, some of which are foundational to the model, and some of which may be adjusted in order to test the hypotheses we have established around inflammasome formation.

First, for simplicity, I have not included caspase-1 in the model, since my aim is to understand the dynamics of inflammasome formation in its early stages. Caspase-1 is recruited later in the inflammasome formation process, and is unlikely to affect the early dynamics of ASC recruitment and cluster formation. Therefore, while it would certainly be useful and interesting to consider the dynamics of caspase-1 recruitment, this is outside the scope of the current model.

I also assume that NLR oligomer formation and ASC aggregation are temporally separated, with oligomer formation occurring before aggregation begins; thus we assume that we have a fixed set of NLR oligomers which we use as ICs for our model. This was a useful assumption, since the mechanisms of NLR oligomer formation are not all well understood; for example, although the structure of the NLRC4 oligomer has been studied in-depth, this is not the case for other NLR oligomers, and furthermore the regulatory processes underlying the formation of these oligomers are complex and have not been fully characterised. Moreover, by not explicitly modelling the oligomer formation process, the models presented here may be applicable to a much wider class of inflammasomes. By taking different distributions of NLR oligomer sizes as ICs for our model, we can test various options for those distributions while avoiding having to make any assumptions about the process of NLR oligomer formation. Furthermore, although it would not be unreasonable to expect that NLR oligomers continue to grow (and NLR-NLR interactions could even be a mechanism by which cluster joining could take place) since we expect the overall abundance of NLRs to be significantly less than the abundance of ASC monomers, it is likely that reactions involving ASC will be

much more important in driving the reaction dynamics. I have therefore disregarded NLR-NLR joining processes for the sake of simplicity.

Since the stoichiometry of NLRs and ASC monomers in the inflammasome is not entirely clear, I have assumed that each NLR can form the basis of a single ASC filament, and that each ASC monomer in a cluster at the end of a filament can recruit at most two further ASC monomers (via a PYD-PYD and CARD-CARD interaction respectively). This is a considerable simplification of the complexity of the protein-protein interactions involved (see for example the comprehensive work of Lu *et al.* and Nambayan *et al.* in documenting the mechanism of ASC-ASC interactions [Lu et al., 2014, Nambayan et al., 2019]). However, this level of detail is sufficient for a mathematical model concerned with the dynamics of ASC depletion and cluster formation over significantly longer timeframes and spatial scales larger than individual reactions between monomers. Furthermore, I have assumed that all the reactions in the model are irreversible, since the emergence of the ASC speck is certainly irreversible, although this does not necessarily imply that on shorter timescales all reactions are entirely irreversible.

A key aspect of any model of inflammasome formation must be a mechanism which prevents aggregation of ASC in the absence of a stimulus. In this model, informed by the prion-like behaviour of ASC, I assume that ASC can only interact directly with NLRs, or with ASC in a cluster. This could be interpreted biologically as ASC gaining an alternative allosteric conformation following interaction with NLRs, which may then be propagated to other ASC monomers, in a manner not dissimilar to NLR oligomerisation. Naturally, there are many other processes at play in the biological system; in particular, I have broadly overlooked the role of post-translational modifications and COPs and POPs. However, the 'allostery-like' description of ASC behaviour is very simple, and does not require the introduction of any more reactants into the system; it also allows us to build on the prion proliferation models of Oosawa and Kasai, and Masel and Nowak [Oosawa and Kasai, 1962, Nowak et al., 1998, Masel et al., 1999]. Furthermore the methods by which PTMs and COPs and POPs regulate inflammasome formation are insufficiently well-documented to allow construction of a mathematical model. It is also worth noting that although the mechanism described above was inspired by the possibility of ASC gaining alternative allosteric conformations, it does not assume that this is in fact the case; the prion-like formulation is simply a simple mathematical model that displays the behaviour required to emulate inflammasome formation, which could

have many biological interpretations.

One feature of inflammasome formation we want to interrogate with this model is whether branching of ASC filaments occurs during inflammasome formation (i.e. initiation of new filaments via CARD-CARD reactions occurring concurrently with filament extension via PYD-PYD interactions), or whether this only occurs via crosslinking once filaments are fully formed. I have therefore allowed filament branching to occur; however, the rate at which this occurs is controlled by a single parameter (c_3) which may be set to zero to remove branching from the model. I have assumed that the rate of branching is proportional to the number of filaments; essentially, I have only allowed branching to occur at the end of filaments. This reflects the fact that the ends of filaments will be more accessible than ASC monomers closer to the NLR oligomer at the centre of a cluster. I have not included crosslinking of existing ASC filaments within clusters, since this is unlikely to affect the early dynamics of ASC clusters, and it would be extremely difficult to keep track of the links within a cluster (although intra-cluster crosslinking would naturally affect the availability of monomers within a cluster available for inter-cluster crosslinking).

Another aspect of conceptual models of the inflammasome which we wish to examine further is the possibility of clusters joining via ASC-ASC interactions. Once again, I have included a parameter which controls the rate of such reactions (c_2) which can be set to zero to exclude joining from the model. I have also assumed that joining of clusters is proportional to the number of filaments in the clusters involved, again because this is an approximation of the number of ASC monomers close to the surface of each cluster.

Finally, I have included the simplest possible spatial structure, dividing the cell into a nuclear and cytosolic region. This is suggested by the structure of the experimental data analysed in Chapter 2. Within these compartments I have assumed that reactants are well-mixed, and ignored finer-grained spatial effects. This is of course a significant simplification, and overlooks the effects of interactions with other cellular structures (such as mitochondria and the cytoskeleton, which have both been implicated in inflammasome formation [Zhou et al., 2011, Misawa et al., 2013, Man et al., 2014a, Li et al., 2017]) as well as molecular crowding. However, these considerations are beyond the scope of this initial inflammasome model. I have also assumed that clusters including NLRs reside only in the cytosolic compartment, since this is where stimuli are initially located. It is also unlikely that clusters will cross the nuclear membrane, both due to their relatively large size (which will inhibit both

their speed of movement and ability to pass through pores in the nuclear membrane), and due to the fact that we rarely observe inflammasomes forming in the nucleus [Cheng et al., 2010].

It is worth noting that accurate measurements of the abundance of ASC and NLRs in the cell, either absolute or relative, are not available. Therefore, it will be necessary to explore a range of possible initial conditions for abundances of ASC monomers, as well as a range of different initial conditions for NLR oligomer sizes and abundances, reflecting the different models of NLR oligomer forms. Values for the reaction rates in this model are likewise not available; however, it is possible to rescale the model to remove the dependence on some variables. Naturally this reduces our ability to fit models to data, although there are too many parameters in the model to be able to carry out straightforward model fits to the data described in Chapter 2.

3.2.3 Expected behaviour of the model

In order to accurately reflect the inflammasome formation process, there are a number of basic properties we wish the model outputs to display.

Since in most cases the inflammasome is characterised by the formation of a single dominant cluster, we expect the total number of clusters to tend towards 1; this would correspond, in the ODE models, to the concentration of clusters tending towards zero. The time at which all initial clusters have merged and a single cluster emerges (or in the continuous model, the time at which the concentration of clusters reaches zero) could therefore be regarded as the time at which an ASC speck has truly formed.

Furthermore, we would expect the population of ASC monomers in the cytosolic and nuclear compartments of the cell to decrease to zero, since inflammasome formation generally depletes all ASC monomers in the cell. This may not coincide with the emergence of the ASC speck (and in fact, in the dataset presented in Chapter 2 we often see the emergence of a speck before cellular ASC is fully depleted), but we would certainly expect ASC depletion in both compartments to be the ultimate fate of the system. The experimental data examined in Chapter 2 also suggests this occurs in a monotonic fashion.

We would therefore expect a successful model of an inflammasome to display these basic properties:

- System ends with a single cluster

- ASC monomers are depleted monotonically throughout the cell, and are ultimately all subsumed into clusters

3.2.4 Full ODE model

A simple mass action formulation of the reactions considered in Table 3.2 gives the equations in System 3.1.

$$\frac{\mathrm{d}\hat{x}_c}{\mathrm{d}t} = \underbrace{-c_1\hat{x}_c\sum_{i,j,k} jx_{i,j,k}}_{\text{C1}} \underbrace{- c_3\hat{x}_c\sum_{i,j,k} ix_{i,j,k}}_{\text{C3}} \underbrace{- p\hat{x}_c\sum_{i,j,k} ix_{i,j,k}}_{\text{P}} + \overbrace{m_1\hat{x}_n}^{\text{M1}} - \overbrace{m_2\hat{x}_c}^{\text{M2}} \tag{3.1a}$$

$$\frac{\mathrm{d}\hat{x}_n}{\mathrm{d}t} = \underbrace{-m_1\hat{x}_n}_{\text{M1}} + \underbrace{m_2\hat{x}_c}_{\text{M2}} \tag{3.1b}$$

For $f < 0, n < 0, a < 0$:

$$x_{f,n,a} = 0 \text{ for all } t \tag{3.1c}$$

And for $f \geq 0, n \geq 0, a \geq 0$:

$$\begin{aligned}\frac{\mathrm{d}x_{f,n,a}}{\mathrm{d}t} = &\underbrace{c_1\hat{x}_c[(n+1)x_{f-1,n+1,a-1} - nx_{f,n,a}]}_{\text{C1}} \\ &\underbrace{+\frac{1}{2}c_2\sum_{i=0}^{f}\sum_{j=0}^{n}\sum_{k=0}^{a} i(f-i)x_{i,j,k}x_{f-i,n-j,a-k} - c_2 f x_{f,n,a}\sum_{i,j,k} ix_{i,j,k}}_{\text{C2}} \\ &\underbrace{+c_3\hat{x}_c[(f-1)x_{f-1,n,a} - fx_{f,n,a}]}_{\text{C3}} \underbrace{+p\hat{x}_c[fx_{f,n,a-1} - fx_{f,n,a}]}_{\text{P}}\end{aligned} \tag{3.1d}$$

Here, $\sum_{i,j,k}$ denotes the triple sum $\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}$.

Note the form of the second term of Equation 3.1d; the factor of $1/2$ is present, in the case of $(i,j,k) \neq (f/2, n/2, a/2)$ since these terms are counted twice in the summation; in the case $(i,j,k) = (f/2, n/2, a/2)$, only one such term is counted in the sum, but the inclusion of the factor of $1/2$ is due to both reactants being functionally identical (see also Section 3.2.1). Note also that the rate of reaction of two clusters of the same size has been slightly adapted from the propensity function given in Table 3.2; since in the ODE model we are considering $x_{f,n,a}$ as concentration rather than total count (for which the propensity functions have been formulated), the $-x_{f,n,a}$ term should be excluded; this gives a good approximation for the joining rate when $x_{f,n,a}$ are sufficiently large.

The following initial conditions are used:

$\hat{x}_c(0) = x_0^{(c)}$

Variable	Expression
X	$\sum_{i,j,k} x_{i,j,k}$
F	$\sum_{i,j,k} i x_{i,j,k}$
N	$\sum_{i,j,k} j x_{i,j,k}$
A	$\sum_{i,j,k} k x_{i,j,k}$
$s_{F,F}$	$\sum_{i,j,k} i^2 x_{i,j,k}$
$s_{N,N}$	$\sum_{i,j,k} j^2 x_{i,j,k}$
$s_{A,A}$	$\sum_{i,j,k} k^2 x_{i,j,k}$
$s_{F,N}$	$\sum_{i,j,k} ij x_{i,j,k}$
$s_{F,A}$	$\sum_{i,j,k} ik x_{i,j,k}$
$s_{N,A}$	$\sum_{i,j,k} jk x_{i,j,k}$

Table 3.3: Variables used in reduced version of ODE model

$\hat{x}_n(0) = x_0^{(n)}$

$x_{f,n,a}(0) = 0$ if $f > 0$ or $a > 0$

$x_{0,n,0}(0)$ are given by the distribution of NLRC4 oligomer sizes; in general, I assume that there is a given initial concentration of clusters X_0, each containing 10 NLRs (although this assumption will be interrogated further in later chapters). In this case, we have:

$$x_{f,n,a} = \begin{cases} X_0 & \text{if } f = 0, n = 10, a = 0 \\ 0 & \text{otherwise} \end{cases}$$

We can reduce the system of equations given in System 3.1 via summation, similarly to the Masel-Nowak model [Masel et al., 1999]; this reduced system of ODEs captures the features of interest while leaving the system more amenable to analysis. The sum variables used are defined in Table 3.3, and equations given in System 3.2. Note that X designates the total concentration of clusters, F denotes the total concentration of filaments, N denotes the total concentration of free NLRs in clusters, and A denotes the total concentration of ASC in clusters.

$$\frac{\mathrm{d}\hat{x}_c}{\mathrm{d}t} = -c_1\hat{x}_c N - c_3\hat{x}_c F - p\hat{x}_c F + m_1\hat{x}_n - m_2\hat{x}_c \tag{3.2a}$$

$$\frac{\mathrm{d}\hat{x}_n}{\mathrm{d}t} = -m_1\hat{x}_n + m_2\hat{x}_c \tag{3.2b}$$

$$\frac{\mathrm{d}X}{\mathrm{d}t} = -\frac{1}{2}c_2 F^2 \tag{3.2c}$$

$$\frac{\mathrm{d}F}{\mathrm{d}t} = c_1\hat{x}_c N + c_3\hat{x}_c F \tag{3.2d}$$

$$\hat{x}_c(0) = \hat{x}_0^{(c)} \qquad s_{F,F}(0) = 0$$
$$\hat{x}_n(0) = \hat{x}_0^{(n)}) \qquad s_{N,N}(0) = 100X_0^2$$
$$X(0) = X_0 \qquad s_{A,A} = 0$$
$$F(0) = 0 \qquad s_{F,N} = 0$$
$$N(0) = 10X_0 \qquad s_{F,A} = 0$$
$$A(0) = 0 \qquad s_{N,A} = 0$$

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -c_1\hat{x}_c N \tag{3.2e}$$
$$\frac{\mathrm{d}A}{\mathrm{d}t} = c_1\hat{x}_c N + c_3\hat{x}_c F + p\hat{x}_c F \tag{3.2f}$$
$$\frac{\mathrm{d}s_{F,F}}{\mathrm{d}t} = c_1\hat{x}_c(2s_{F,N} + N) + c_2 s_{F,F}^2 + c_3\hat{x}_c(2s_{F,F} + F) \tag{3.2g}$$
$$\frac{\mathrm{d}s_{N,N}}{\mathrm{d}t} = c_1\hat{x}_c(N - 2s_{N,N}) + c_2 s_{F,N}^2 \tag{3.2h}$$
$$\frac{\mathrm{d}s_{A,A}}{\mathrm{d}t} = c_1\hat{x}_c(2s_{N,A} + N) + c_2 s_{F,A}^2 + (c_3 + p)\hat{x}_c(2s_{F,A} + F) \tag{3.2i}$$
$$\frac{\mathrm{d}s_{F,N}}{\mathrm{d}t} = c_1\hat{x}_c(s_{F,F} - s_{F,N} - N) + c_2 s_{F,F}s_{F,N} + c_3\hat{x}_c s_{F,N} \tag{3.2j}$$
$$\frac{\mathrm{d}s_{F,A}}{\mathrm{d}t} = c_1\hat{x}_c(s_{N,A} + s_{F,N} + N) + c_2 s_{F,F}s_{F,A} + c_3\hat{x}_c(s_{F,F} + s_{F,A} + F) + p\hat{x}_c s_{F,F} \tag{3.2k}$$
$$\frac{\mathrm{d}s_{N,A}}{\mathrm{d}t} = c_1\hat{x}_c(s_{N,N} - s_{N,A} - N) + c_2 s_{F,N}s_{F,A} + (c_3 + p)\hat{x}_c s_{F,N} \tag{3.2l}$$

The intial conditions for this system of equations are as follows:

Note that $\frac{\mathrm{d}A}{\mathrm{d}t} + \frac{\mathrm{d}\hat{x}_c}{\mathrm{d}t} + \frac{\mathrm{d}\hat{x}_n}{\mathrm{d}t} = 0$, so from the ICs we have the conservation relation $\hat{x}_c + \hat{x}_n + A = \hat{x}_0^{(c)} + \hat{x}_0^{(n)}$.

Interestingly, the variables $\hat{x}_c, \hat{x}_n, F$ and N do not depend on any other variables, but I have included other variables for interest, to create a larger closed system.

3.2.5 Stochastic model

As will be seen in later chapters, we encounter issues with the ODE model as we reach situations with few clusters; the use of continuous variables representing concentrations is then problematic, and furthermore our approximation of joining rates (in which the possibility of self-joining of clusters is ignored) no longer holds when there are few clusters. These considerations are particularly important when considering events which occur late in the inflammasome formation process (for example, the emergence of a single ASC speck). These issues can be resolved by using a stochastic version of the model.

The reactions outlined for the ODE model can easily be used in a stochastic context instead, using the propensity functions given in Table 3.2. Realisations of this model can then be simulated using, for example, the Gillespie algorithm [Gillespie, 1972]. It is also possible to use more analytical techniqes, for example, using a master equation approach, or applying a CTMC approach as used by Yvinec *et al.* in stochastic BD models, D'Orsogna *et al.* in ML models, and Kashchiev in NP models [Yvinec et al., 2012, Yvinec et al., 2016, D'Orsogna et al., 2015, Kashchiev, 2018]. However, due to the extremely large number of reactants involved, this system is somewhat intractable to such analysis. Furthermore, difficulties arise because of the non-constant 'kernel' of cluster joining, the concurrent processes of cluster growth and joining, and movement between cellular compartments. This motivates the construction of a simplified model which is more amenable to analysis.

3.3 Simplified model

I have also created a simplified version of the inflammasome formation model which is much more analytically tractable; in its ODE form it can be solved analytically, and the stochastic form is also considerably more amenable to analysis. This analytical approach is useful since it gives much more precise and easily-interpretable solutions than the numerical integration and simulation techniques which are necessary for analysing the full model. However, naturally there may be disadvantages to removing key elements in the full model, which can be seen by comparing results for the full model and simplified model.

3.3.1 Description of simplified model

The alterations to the full model which give the simplified model are as follows:

- Assume cluster merging rate is independent of cluster size
- Consider the total population of ASC monomers in the cell rather than separating into subpopulations based on subcellular location
- Assume free NLRs and filaments behave identically in terms of ASC recruitment

Removing the size dependence of cluster joining reactions may be somewhat biologically implausible, but this simplifies considerably the coagulation aspect of the model by essentially setting the joining kernel to be constant. Furthermore, it is useful to remove the size

dependence of joining to establish what role, if any, it plays in determining the behaviour of the model.

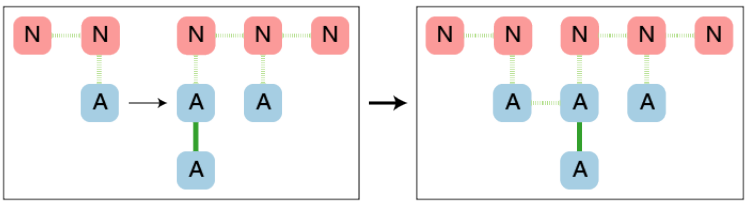
Combining the two cellular compartments to give a single population of ASC monomers means we lose some of the structure present in the experimental data — we cannot compare the output of this model to the separate nuclear and cytosolic traces derived in Chapter 2, but instead must consider the cell as a whole. However, once again, this simplifies the mathematics considerably (especially since we expect movement between compartments to happen on a much faster timescale than cluster growth and joining). Furthermore, ASC monomers in the nuclear compartment are essentially passive in the full model; it is therefore useful to establish whether the nuclear compartment really displays any mathematically interesting behaviour, or if its inclusion in the model merely adds a lag to the depletion of ASC monomers.

Finally, the conflation of filaments and free NLRs once again is a considerable deviation from our biological assumptions; in developing the full model, it was important to keep track of both of these variables since ASC joining to NLRs and to other ASCs cannot be assumed to happen at the same rate. However, conflating filaments and free NLRs allows us to use only two subscripts to denote cluster size. This simplification may be justified by the assumption that the initial number of NLRs is small compared to the initial number of ASC monomers; therefore, unless the rate of joining of ASC to NLRs is extremely slow, the population of free NLRs will be depleted relatively quickly compared to the timeframe of inflammasome formation and ASC monomer depletion. Furthermore, in the case of the NLRC4 inflammasome, both ASC binding with NLRC4 and putative ASC filament branching reactions occur via CARD-CARD interactions; this suggests that a similar rate for both reactions is not unreasonable.

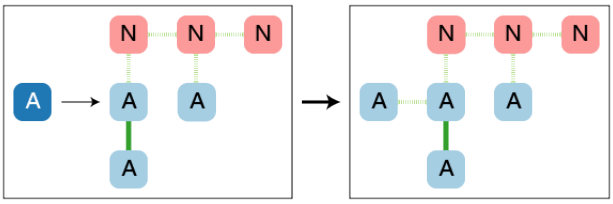
The variables and reactions used in the simplified model are very similar to those in the full model; these are laid out in Figure 3.3 and Tables 3.4 and 3.5.

3.3.2 Simplified ODE model

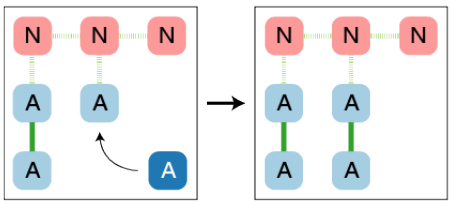
The reactions described in Table 3.5 can be expressed using the following ODEs:



(a) **Reaction C2:** Two clusters join via CARD-CARD interactions between active ASC in a cluster. A new cluster is formed, where the number of free NLRC4 and filaments is the sum of those of the constituent clusters. Here, $x_{1,1,1}+x_{2,1,3} \rightarrow x_{3,2,4}$. The rate of clusters joining is independent of cluster size.



(b) **Reaction C:** An inactive ASC monomer joins a cluster via a CARD-CARD interaction with an NLR or active ASC in a cluster, and is activated. The number of filaments in the cluster increases by one. Here, $\hat{x}_c + x_{2,1,3} \rightarrow x_{3,1,4}$.



(c) **Reaction P:** An inactive ASC monomer joins a cluster via PYD-PYD interaction with active ASC in a cluster, and is activated. The filament the monomer joins is extended, but the total number of free NLRC4 and filaments in the cluster is unchanged. Here, $x_{2,1,3} + \hat{x}_c \rightarrow x_{2,1,4}$

Figure 3.3: Reactions considered in the simplified inflammasome model

Variable	Explanation
$\hat{x}$	Total concentration of ASC monomers present throughout the cell
$x_{f,a}$	Total concentration of clusters in cytosolic compartment with f growing ASC filaments and NLR monomers, and a ASC monomers

Table 3.4: Variables used in ODE and stochastic versions of simplified model

Reference	Reaction	Propensity function
C	$\hat{x} + x_{f,a} \xrightarrow{cf} x_{f+1,a+1}$	$cf\hat{x}x_{f,a}$
C2	$x_{f,a} + x_{i,j} \xrightarrow{c_2 fi} x_{f+i,a+j}$	$c_2 fix_{f,a}x_{i,j}$ for $(f,a) \neq (i,j)$, $\frac{1}{2}c_2 f^2 x_{f,a}(x_{f,a}-1)$ otherwise
P	$\hat{x} + x_{f,a} \xrightarrow{pf} x_{f,a+1}$	$pf\hat{x}x_{f,a}$

Table 3.5: Rates of reactions considered in ODE and stochastic versions of simplified model.

Variable	Expression
X	$\sum_{i,j,k} x_{i,j,k}$
F	$\sum_{i,j,k} ix_{i,j,k}$

Table 3.6: Variables used in reduced version of simplified ODE model

$$\frac{\mathrm{d}x_{f,a}}{\mathrm{d}t} = \underbrace{\frac{1}{2}c_2 \sum_{i=0}^{f}\sum_{j=0}^{a} x_{f-i,a-j}x_{i,j} - c_2 x_{f,a}\sum_{i,j} x_{i,j}}_{\text{C2}} + \underbrace{c\hat{x}[(f-1)x_{f-1,a-1} - fx_{f,a}]}_{\text{C}} + \underbrace{p\hat{x}[fx_{f,a-1} - fx_{f,a}]}_{\text{P}} \tag{3.3a}$$

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = \underbrace{-(c+p)\hat{x}\sum_{i,j} ix_{i,j}}_{\text{C,P}} \tag{3.3b}$$

We assume that initially there is a concentration of X_0 NLR oligomers each containing 10 NLRs; this gives initial conditions

$$x_{f,a} = \begin{cases} X_0 & \text{if } f = 10, a = 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.4a}$$

$$\hat{x}(0) = \hat{x}_0 \tag{3.4b}$$

Similarly to the full model (See System 3.2), the ODEs given in System 3.3 can be reduced to a system involving 'summation variables':

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = -(c+p)\hat{x}F \tag{3.5a}$$

$$\frac{\mathrm{d}X}{\mathrm{d}t} = -\frac{1}{2}c_2 X^2 \tag{3.5b}$$

$$\frac{\mathrm{d}F}{\mathrm{d}t} = c\hat{x}F \tag{3.5c}$$

The initial conditions are as follows:

$$\hat{x}(0) = \hat{x}_0 \tag{3.6a}$$
$$X(0) = X_0 \tag{3.6b}$$
$$F(0) = 10X_0 \tag{3.6c}$$

This system is very similar in form to the reduced system for the full model (Equation 3.2); note in particular the form of the equation for X, in which the right-hand side is now dependent only on X, due to the constant joining rate.

3.3.3 Simplified stochastic model

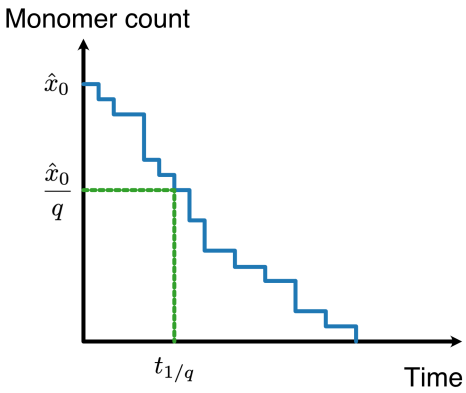
As with the full model, it is useful to have a stochastic version of the simplified model, especially when considering events which occur late in the inflammasome formation process. We can create a simplified stochastic model of inflammasome formation using the reactions and propensity functions described in Table 3.5, and simulate results using, for example, the Gillespie algorithm [Gillespie, 1972]. One primary motivation for constructing the simplified model is that this stochastic model is sufficiently simple that it can be analysed analytically, without the need to rely on simulations. We can therefore consider the reactions within the framework of a CTMC, for which there are numerous existing mathematical tools for analysis.

3.4 Characteristic times

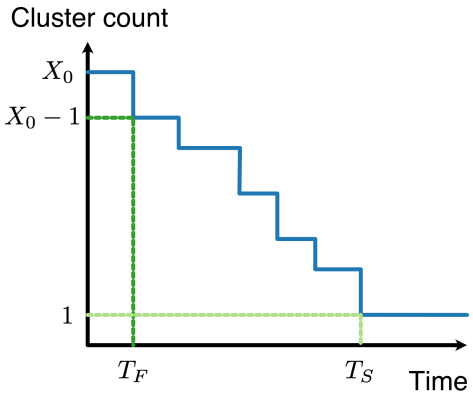
In order to compare results from the full model, simplified model and the dataset presented in Chapter 2, my main approach will be the use of characteristic times. These have been defined for the experimental data in the previous chapter (see Sections 2.3.3 and 2.3.4); I will now describe the characteristic times for the models presented in this chapter, and explain how they link to the characteristic times derived for the experimental data. The characteristic times are summarised in Figure 3.4.

3.4.1 First joining time

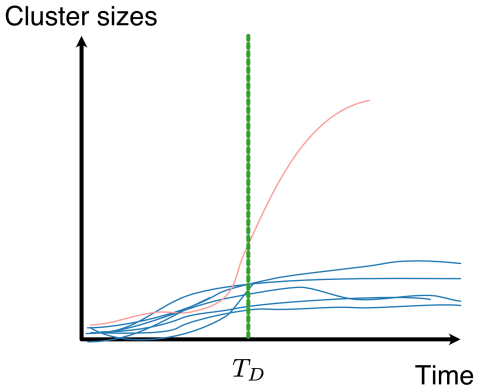
The time at which two clusters first join is a useful marker of the transition from cluster growth by monomer addition only to a phase of growth in which cluster joining begins. This quantity cannot be measured in the experimental data presented in Chapter 2, but is useful for understanding the differences between the simplified and full models, and also in how



(a) Monomer depletion characteristic times $t_{1/q}$: time for monomer count to decrease by a fraction $1/q$. A possible example of monomer count over time is shown in blue; the time taken for monomer count to decrease by a fraction $1/q$ is shown in green.



(b) Single cluster emergence time T_S: time at which only one cluster remains. First joining time T_F: time of first joining event. A possible example of total cluster count over time is shown in blue. Time until first joining event is shown in dark green. Time until only one cluster remains is shown in light green.



(c) Dominant cluster emergence time T_D: time at which one cluster becomes much larger than other clusters at that time. Lines in blue and pink represent the sizes of all clusters over time for one run of the model; the time at which the cluster which eventually dominates (shown in pink) first diverges in size from the other clusters is shown in green.

Figure 3.4: Schematics illustrating characteristic times for the inflammasome models; all figures show one possible scenario for monomer depletion and cluster growth, for a single run of the model.

different parametrisations of each model affect the kinetics of the inflammasome formation process. I will refer to this characteristic time as the 'first joining time', or T_F, throughout this thesis.

3.4.2 Inflammasome emergence time

The time at which the inflammasome is fully formed is naturally extremely important in this analysis, and can be compared to the time at which the ASC speck appears in the experimental data; in practice, this is taken to be the second speck trace breakpoint (see Section 2.3.3), although the initial lag before inflammasome formation begins must be subtracted first. A natural comparison can be drawn to the earliest time in the models at which all clusters have merged into a single cluster; I will refer to this as the single cluster emergence time, or T_S, throughout. There may still be unbound ASC monomers present at this time, but all clusters of size greater than 1 will have merged into a single cluster, which may continue to grow via recruitment of unbound ASC. Note that this quantity cannot be calculated for the ODE variants of the models, due to the continuous nature of the variables. However, another possible characterisation of the time of inflammasome formation is the time at which a single large cluster has emerged; as we will see, this is characteristic of many solutions for both models, and is analogous to the emergence of a giant component in random networks, or indeed to gelation in Smoluchowski and ML models, although without the mass loss characteristics. This could be a better characterisation of the time at which a large ASC speck emerges. I will refer to this characteristic time as the dominant cluster time, or T_D.

I have defined the dominant cluster emergence time as the time at which one cluster becomes an outlier in terms of the size distribution of all other clusters at that time. I have used a simple measure to calculate the threshold for a cluster becoming a 'large outlier' based on the mean and standard deviation as follows:

$$\text{Threshold} = \text{mean} + 3 \times \text{standard deviation} \tag{3.7}$$

If the cluster sizes are normally distributed, we would expect around 0.1% of cluster sizes to lie above this value.

3.4.3 Monomer depletion characteristic times

As described previously, the time of single cluster emergence generally does not coincide with the time of single cluster emergence. It is therefore also useful to designate characteristic times which describe monomer depletion, for both cellular compartments in the case of the full model. To quantify the rate of monomer depletion, I will use the monomer depletion characteristic times $t_{1/q}$, the time taken for the monomer population in a compartment (or throughout the cell, in the case of the simplified model) to decrease by a fraction $1/q$. These can be compared directly to $t_{1/q}$ as calculated for the experimental data (see Section 2.3.4), although once again the initial lag should be subtracted first.

3.5 Discussion

3.5.1 Summary

In this chapter, I have presented two models of inflammasome formation; a 'full' model, which includes the key biological aspects of inflammasome formation, and a simplified model, which is an adapted version of the full model which is more amenable to mathematical analysis. Both models have the potential to explore unanswered questions in inflammasome formation; in particular, both allow branching of ASC filaments and joining of clusters to be included or excluded, and allow for the flexible input of different distributions of NLR oligomer sizes as ICs. I have also established the basic expectations we have for solutions of a model of inflammasome formation; namely, that the system approaches a state with a single inflammasome, via monotonic depletion of ASC monomers. Finally, I have defined the characteristic times which will be used to analyse the models, and compared them to the experimental data introduced in Chapter 2.

Both the full and simplified models will be explored further in the following chapters, and the differences between solutions to these models will be explained.

3.5.2 Further work

The models I have presented here are only a starting point for the possibilities of modelling inflammasome formation. While I have established a simple foundation for future modelling attempts, these models immediately prompt a number of natural extensions.

Firstly, other reactants could be included in the model. Adding caspase-1 would be useful, to interrogate the assumption that caspase-1 does not affect the dynamics of inflammasome formation. Furthermore, it is still not clearly understood why the inflammasome structure is required for caspase-1 recruitment at all; studying the dynamics of caspase-1 recruitment using a model similar to those presented might shed some light on this issue. However, more quantitative data describing the total abundance of processed caspase-1 over time would be extremely useful for developing such a model. It would also be a natural extension to add PTMs, COPs or POPs to the model. This could be as simple as introducing a process by which NLRs are phosphorylated or dephosphorylated, with the two NLR populations interacting differently with ASC. However, it would be useful to have more information available about the processes through which PTMs occur before including this factor in the model.

It would also be possible to add reversibility of reactions into the model; in particular, this may allow an alternative characterisation of prevention of ASC clustering in the absence of stimuli (for example, it would be possible to create a model in which ASC monomers may form clusters in the absence of stimuli, but the rate of dissociation is very high unless the clusters are in some manner stablised by interaction with an activated NLR). This could be a useful change to the current simple model of prion-like ASC activation, since the model currently does not account for spontaneous ASC speck formation in ASC overexpression systems.

A more sophisticated treatment of NLR oligomerisation would be another natural step in improving the existing models; there is nothing to suggest that NLR reactions cannot continue once ASC recruitment has begun. However, once again it may be easier to create such a model if we have more experimental data on NLR behaviour over time. If we were, for example, to include NLR oligomerisation as an element in a model of inflammasome formation, we would require much more information about the input of stimulus into the cell over time.

One interesting extension to the existing inflammasome model would be to examine the structure of the final inflammasome more closely. Currently, we do not track the precise structure of the inflammasome over time but only the the number of free NLRs, ASC filaments and total ASC monomers in the inflammasome. Therefore, it is difficult to assess the degree of crosslinking or branching in the completed inflammasome. A natural strategy to approach

such a problem would be to consider the inflammasome as a network, with nodes representing individual proteins and edges representing reactions between proteins. It would then be much easier to include crosslinking reactions within a cluster. Furthermore, we would then have a network representing the complete inflammasome which could be analysed for properties such as connectedness. This could allow us to investigate the relative importance of different protein interactions in inflammasome formation, and investigate the appearance of 'Medusa's head' phenotypes in ASC overexpression systems, which could correspond to low levels of connectivity in inflammasome networks.

The simple two-compartment structure of the full model also prompts the development of a model inflammasome formation with finer spatial granularity. The reactions in either ODE model could easily be extended to PDEs including diffusion. This would allow us to examine in more detail the role of spatial effects in inflammasome formation. In particular, the models presented here assume a well-mixed environment, when clearly during inflammasome formation we approach a situation in which ASC is highly concentrated in one area of the cell. Extending the models presented here to fully spatial models would allow the effects of molecular crowding to be considered more easily, as well as introducing interactions with other cellular structures. It would be particularly interesting to consider interactions with elements of the cytoskeleton; in particular, it is unclear to what extent interactions with actin or microtubules might focus the movement of ASC towards the final ASC speck.

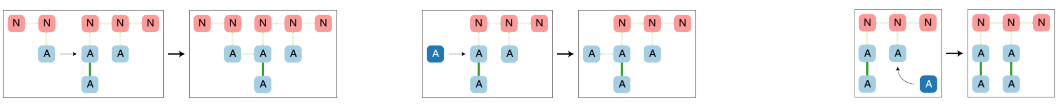
Finally, a recurring theme in this modelling process has been the need for more experimental data. This is the case in a number of different contexts; first of all, in order to develop the characterisation of some processes in more depth (for example, the formation of NLR oligomers), it would be helpful to have a clearer understanding of, and ideally more qualitative data describing, how these processes occur in the biological context. Secondly, in order to allow development of models with, for example, a spatial element, more high-resolution, three dimensional datasets would be extremely useful. Finally, as we will see in later chapters, the lack of concentration data and reaction rate parameters will be a hindrance in attempts to fit these models to data; however, hopefully the creation of these models will clearly motivate the need for acquisition of such data, and provide a clear mathematical and notational framework for deriving rate parameter values.

Chapter 4

Mathematical analysis of the simplified model of inflammasome formation

4.1 Introduction

In this chapter I will present analysis of both the deterministic continuous ODE and the discrete stochastic formulations of the simplified inflammasome model (Figure 4.1). The benefit of the simplified model is that it is considerably more amenable to mathematical approaches than the full model; unusually for many mathematical models of biological processes, it is possible to derive analytical solutions both for the simplified ODE model and for the distributions of characteristic times for the stochastic model. These analytical solutions are especially useful, in the absence of measured parameter values, for establishing the behaviour of the model over wide ranges of parameter values without the need to run large numbers of simulations, and for fitting distributions of characteristic times to experimental data (this will be carried out in Chapter 6).



(a) Reaction C2 (Cluster joining)

(b) Reaction C (Filament branching)

(c) Reaction P (Filament extension)

Figure 4.1: The simplified inflammasome model revisited

I will begin by providing an analytical solution to the reduced ODE system presented in the preceding chapter (System 3.5), which can then be used along with a generating function approach to solve the simplified ODE model analytically (3.3). I will use these results to examine the distribution of cluster sizes over time — in particular, using asymptotic techniques to establish the behaviour of the model at small and large times, and for small and large parameter values. This approach is useful for establishing the role of each process in inflammasome formation, especially the relative roles of ASC filament extension and branching. However, I will also demonstrate that the model does not exhibit the required behaviour at large times, due to the fact that in this model ASC monomers are never fully depleted.

I will then present analysis of the stochastic model, including analytical expressions for the distribution of the characteristic monomer depletion times, single cluster emergence time, and first joining time. These can be directly compared to data (see Chapter 6). I will also demonstrate the role of different processes and parameters in determining distributions of characteristic times. A particularly notable result is that the initial number of NLR oligomers plays a comparatively small role in determining the time of single cluster emergence in the simplified model.

4.2 Analysis of the ODE model

4.2.1 Full solution for the simplified ODE model

In order to establish the concentrations of ASC monomers and clusters over time, I begin by solving the reduced system first presented in Chapter 3:

$$\begin{aligned} \frac{\mathrm{d}\hat{x}}{\mathrm{d}t} &= -(c+p)\hat{x}F \\ \frac{\mathrm{d}X}{\mathrm{d}t} &= -\frac{1}{2}c_2X^2 \\ \frac{\mathrm{d}F}{\mathrm{d}t} &= c\hat{x}F \end{aligned} \qquad \text{(3.5 revisted)}$$

The equation for X does not involve any other variables and is separable. The system for $\hat{x}$ and F can be solved by noting that $c\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} + (c+p)\frac{\mathrm{d}F}{\mathrm{d}t} = 0$. Therefore, the quantity $N = c\hat{x} + (c+p)F$ is conserved. From our ICs, $N = c\hat{x}_0 + (c+p)X_0N_0$. Thus, System 3.5 can be solved by, for example, substituting $F = \frac{N-c\hat{x}}{c+p}$ into the equation for $\hat{x}$. We therefore arrive at the following solutions:

$$\hat{x}(t) = \frac{N\hat{x}_0}{c\hat{x}_0 + (c+p)X_0N_0e^{Nt}} \tag{4.1a}$$

$$X(t) = \frac{2X_0}{c_2X_0t + 2} \tag{4.1b}$$

$$F(t) = \frac{NX_0N_0e^{Nt}}{c\hat{x}_0 + (c+p)X_0N_0e^{Nt}} \tag{4.1c}$$

A full analytical solution can be found for the full ODE system given in System 3.3 using the generating function $G(y,z,t) = \sum_{f,a} x_{f,a}(t)y^f z^a$. This is similar to the approach used by Davies *et al.* to solve the Smoluchowski equations, outlined in Section 1.5.1 and [Davies et al., 1999].

$$\begin{aligned}\frac{\mathrm{d}x_{f,a}}{\mathrm{d}t} &= \frac{1}{2}c_2\sum_{i=0}^{f}\sum_{j=0}^{a} x_{f-i,a-j}x_{i,j} - c_2x_{f,a}\sum_{i,j}x_{i,j} + c\hat{x}[(f-1)x_{f-1,a-1} - fx_{f,a}] \\ &\quad + p\hat{x}[fx_{f,a-1} - fx_{f,a}] \\ \frac{\mathrm{d}\hat{x}}{\mathrm{d}t} &= -(c+p)\hat{x}\sum_{i,j} ix_{i,j}\end{aligned}$$

(3.3 revisited)

Substituting G into Equation 3.3a (shifting summations where necessary) gives the following PDE:

$$\frac{\partial G}{\partial t} = \frac{1}{2}c_2G^2 - c_2XG + \hat{x}y(c(yz-1) + p(z-1))\frac{\partial G}{\partial y} \tag{4.2}$$

The initial conditions are given by $G(y,z,0) = X_0y^{N_0}$.

We can now substitute the values of X and $\hat{x}$ derived previously (System 4.1) to give a PDE for G:

$$\frac{\partial G}{\partial t} = \frac{1}{2}c_2G^2 - \frac{2c_2X_0}{c_2X_0t+2}G + \frac{N\hat{x}_0}{c\hat{x}_0 + 10(c+p)X_0e^{Nt}}y(c(yz-1) + p(z-1))\frac{\partial G}{\partial y} \tag{4.3}$$

Equation 4.3 can now be solved using the method of characteristics. The characteristics on paths parametrised by s are given by the following system of ODEs:

$$\frac{\mathrm{d}t}{\mathrm{d}s} = 1 \tag{4.4a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}s} = -\hat{x}y(c(yz-1) + p(z-1)) \tag{4.4b}$$

$$\frac{\mathrm{d}z}{\mathrm{d}s} = 0 \tag{4.4c}$$

$$\frac{\mathrm{d}G}{\mathrm{d}s} = \frac{1}{2}c_2G^2 - c_2X(t)G \tag{4.4d}$$

If we parametrise intial conditions on $s = 0$ as $t = 0, y = \xi, z = \eta$, and $G = X_0\xi^{N_0}$, clearly $t = s$ and $z = \eta$. After substituting $t = s$ (noting that Equation 4.4b is separable and Equation 4.4d is a Bernoulli equation) the system of equations can be solved to derive a general solution for G:

$$G = \frac{4X_0}{(c_2X_0t + 2)^2 g(y, z, t)^{N_0} - c_2X_0t(c_2X_0t + 2)} \tag{4.5a}$$

$$g(y, z, t) = f_2(t))^{f_1(z)} \left(\frac{z}{f_1(z)} + \frac{1}{y}\right) - \frac{z}{f_1(z)} \tag{4.5b}$$

$$f_1(\eta) = \frac{p(\eta - 1) - c}{c} \tag{4.5c}$$

$$f_2(t) = \frac{c\hat{x}_0e^{-Nt} + (c + p)X_0N_0}{N} \tag{4.5d}$$

The values of $x_{f,a}$ can be derived by finding the coefficients of powers of y and z in a power series expansion of G.

This analytical expression for the double expansion in y and z does not have a simple closed form. However, the distributions of numbers of filaments and ASC monomers in each cluster can also be derived computationally at each time point by calculating the Taylor expansion for G in terms of the variables y and z (Figure 4.2). The mean and variance of the number of filament and ASC monomers in each cluster increases over time; the skewness of the distribution of cluster sizes (either by filament or ASC monomer count) also increases over time. However, it is simpler to proceed by considering the moments of the distributions of cluster sizes over time.

4.2.2 Moments and asymptotic analysis

The moments of the distribution of cluster sizes are defined as follows:

$$M_{p,q} = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} j^p k^q x_{j,k}(t) \tag{4.6}$$

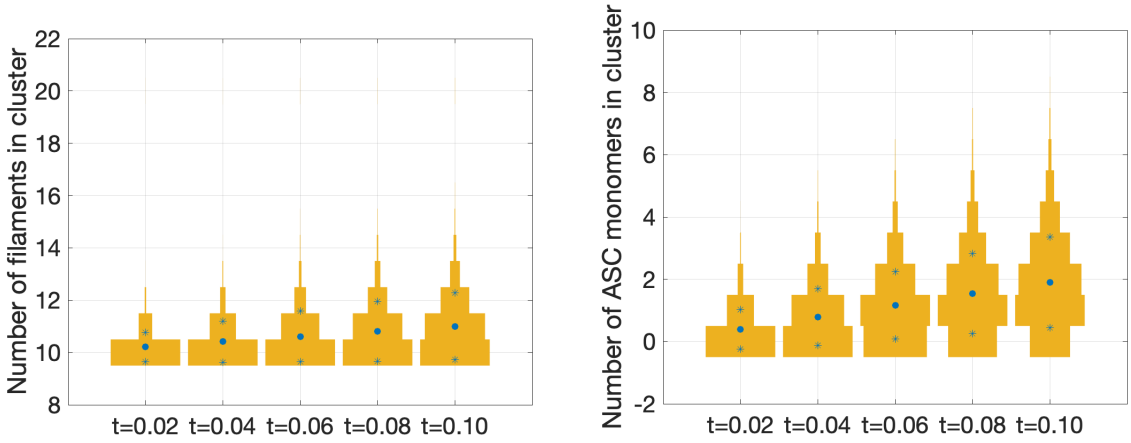
Note that $M_{0,0}$ and $M_{1,0}$ correspond to X and F respectively.

The moments can be derived from the generating function:

$$M_{p,q} = \left[\left(\frac{\partial}{\partial y}\right)^p \left(\frac{\partial}{\partial z}\right)^q G(e^y, e^z, t)\right]_{x=0,y=0} \tag{4.7}$$

$$M_{0,1} = N_0X_0 \left(\frac{p + c}{c}\right)\left(\frac{1}{f_2(t)} - 1\right) \tag{4.8a}$$

$$M_{2,0} = \frac{N_0X_0}{f_2(t)} \left(\frac{c_2N_0X_0t + N_0 + 1}{f_2(t)} - 1\right) \tag{4.8b}$$



(a) Proportions of number of filaments per cluster

(b) Proportions of number of ASC monomers per cluster

Figure 4.2: Violin plots showing proportion of number of filaments and ASC monomers in each cluster over time, calculated using the Taylor expansion of the generating function for cluster size. The width of each bar corresponds to the proportion of clusters with size corresponding to the y-value. Calculated mean size of cluster at each timepoint is shown with a blue circle, and mean $\pm$ standard deviation with a blue star. Distributions are calculated for $p = c_2 = c_3 = 1, X_0 = 0.1, \hat{x}_0 = 1, N_0 = 10$. Both mean filament and ASC monomer counts increase with time; both distributions show considerable positive skewness, with skewness and spread increasing over time.

$$M_{0,2} = N_0 X_0 \left(\frac{p+c}{c}\right)\left[\left(\frac{c+p}{c}\right)\left(\frac{1}{f_2(t)}-1\right)^2 (N_0 X_0 c_2 t + N_0 + 1) + \left(\frac{c+2p}{c}\right)\left(\frac{1}{f_2(t)}-1\right) + \frac{2p\log(f_2(t))}{c f_2(t)}\right] \tag{4.8c}$$

$$M_{1,1} = \frac{N_0 X_0}{f_2(t)}\left[\left(\frac{c+p}{c}\right)\left(\frac{1}{f_2(t)}-1\right)(N_0 X_0 c_2 t + N_0 + 1) + \frac{p\log(f_2(t))}{c f_2(t)}\right] \tag{4.8d}$$

The moments can be used to calculate the centralised moments, including the mean number of filaments per cluster:

$$\begin{aligned}\mu_F &= \frac{M_{1,0}}{M_{0,0}} \\ &= \frac{N_0(c_2 X_0 t + 2)}{2 f_2(t)}\end{aligned} \tag{4.9}$$

And the variance of the number of filaments per cluster:

$$\begin{aligned}\sigma_F^2 &= \frac{M_{2,0}}{M_{0,0}} - \mu_f^2 \\ &= \frac{N_0(c_2 X_0 t + 2)}{2 f_2(t)}\left(\frac{1}{f_2(t)}\left(\frac{c_2 N_0 X_0 t}{2} + 1\right) - 1\right)\end{aligned} \tag{4.10}$$

The mean number of ASC monomers per cluster is

$$\begin{aligned}\mu_A &= \frac{M_{0,1}}{M_{0,0}} \\ &= \frac{N_0(c+p)(c_2X_0t+2)}{2c}\left(\frac{1}{f_2(t)}-1\right)\end{aligned} \tag{4.11}$$

and the variance of the number of ASC monomers per cluster is

$$\begin{aligned}\sigma_A^2 &= \frac{M_{0,2}}{M_{0,0}} - \mu_A^2 \\ &= N_0X_0\left(\frac{p+c}{c}\right)\left[\left(\frac{c_3+p}{c}\right)\left(\frac{1}{f_2(t)}-1\right)^2\left(\frac{N_0X_0c_2t}{2}+1\right)+\left(\frac{c+2p}{c}\right)\left(\frac{1}{f_2(t)}-1\right)\right. \\ &\qquad \left. + \frac{2p\log(f_2(t))}{cf_2(t)}\right]\end{aligned} \tag{4.12}$$

We can use these equations to gain insight into the behaviour of distributions of clusters over time. In particular, we can see that the mean and variance of the number of both filaments and ASC monomers in each cluster increases over time, which is in agreement with the distributions calculated in the previous section (Figure 4.2). We can also use asymptotic solutions to understand the behaviour of solutions at early and late times, and for extreme parameter values.

Let us first assume that all parameters are $O(1)$. The asymptotic expansions for large and small times are given in Table 4.1, and plots of $\mu_F, \mu_A, \sigma_F^2, \sigma_A^2$ and $\hat{x}$ are given in Figures 4.3 - 4.7.

At small times, $\hat{x}$ and X are approximately linearly decreasing, σ_F^2, μ_A and σ_A^2 are linearly increasing from zero, and μ_F is linearly increasing from N_0. The rate of decrease in $\hat{x}$ is proportional to $\hat{x}_0, X_0$ and N_0, while the rate of decrease in X is proportional to c_2 and X_0; thus, at early times, the rate of decrease of monomers and clusters is determined only by the abundance of monomers and filaments, or the number of clusters and joining rate respectively. The rate of increase of μ_A and σ_A^2 are proportional to $N_0, \hat{x}_0$ and $c+p$, while μ_F and σ_F^2 increase at a rate proportional to $N_0c(\hat{x}_0 + N_0X_0/2)$. Thus, we see faster cluster growth in terms of number of ASC monomers in the cluster when both p and c are larger, whereas naturally the filament growth rate is dependent on the value of c but not p. The variance of cluster size (both by number of filaments and number of ASC monomers) increases over time as cluster sizes initially diverge from the initial conditions, in which all clusters contain 10 NLRs. The initial increases for both mean and variance of cluster sizes are especially

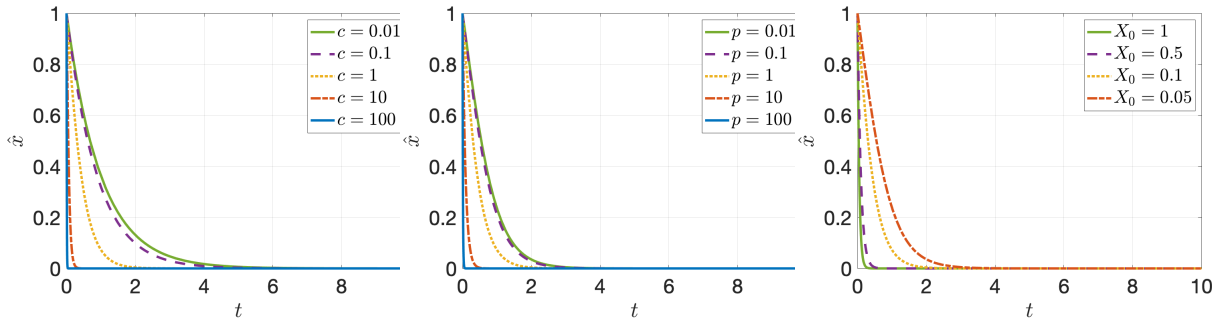
steep when c is large, since when branching reactions occur, this creates further filaments, which increases the rate of both C and P-type reactions; in contrast, increasing p increases the growth rate of clusters, but we do not observe a positive feedback loop as for increasing c, since no new filaments are formed.

Notably, the solutions for μ_F, μ_A, σ_F^2 and σ_A^2 at early times approximate the solutions of the same system with $\hat{x} = \hat{x}_0$ held constant. It is also notable that there is no initial 'lag' either in monomer depletion or cluster growth; in fact, for mean and variance of cluster sizes, for all parametrisations shown here there is generally an initial sharp increase which slows at later times. We will see in the next chapter that such a lag can occur for the full model; this is a key difference between the behaviour of the simplified and full models.

At large times, $\hat{x}$ and X tend towards zero as desired, as $O(e^{-Nt})$ and $O(t^{-1})$ respectively. μ_f and μ_a increase approximately linearly for large t, as the system becomes dominated by lower concentrations of large clusters. However, σ_A^2 increases linearly, and σ_F^2 increases quadratically with time. This violates our conditions on successful modelling of the inflammasome laid out in Section 3.2.3; if the system tends towards a steady state with a single cluster as experimental observations show, we would expect the variance of the system to decrease at large times until it reaches zero. Furthermore, we would not expect the mean size of the cluster to continue increasing indefinitely; as the ASC monomers in the system are depleted, we would expect the final ASC speck to approach a maximum size, when it contains all ASC originally in the cell. This failure of the simple ODE model to converge to a mean cluster size is a result of the fact that the concentration of monomers is continuous; thus the nonzero monomer concentration can continue contributing to cluster growth indefinitely.

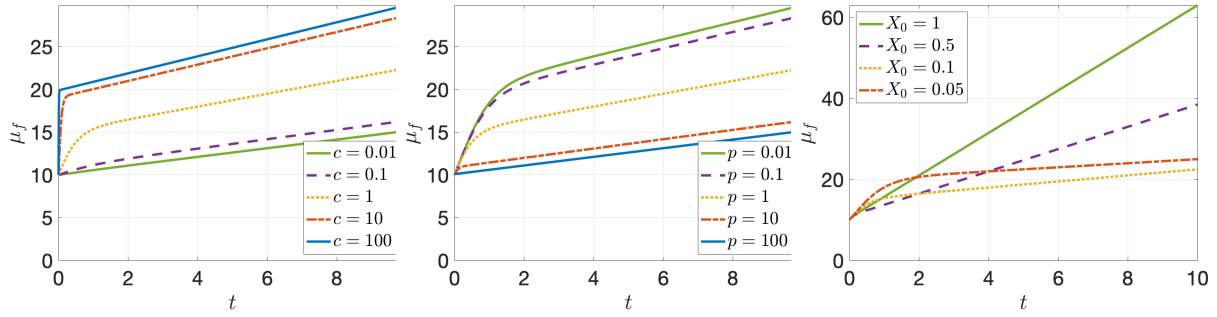
4.2.3 Models without branching or with branching only

In order to clarify the roles played by the branching (C) and filament extension (P) reactions in the model, I will now consider the solutions of the model when each of these reactions is removed in turn; this corresponds to setting $c = 0$ or $p = 0$. The case $c = 0$ corresponds to a system in which clusters grow only via filament extension; the case $p = 0$ corresponds to a system in which clusters grow only by reactions which seed new filaments. Consideration of these extreme systems will allow us to shed light on how filament-extension-like and branch-like growth affect the distributions of cluster growth over time. Furthermore, this will enable us



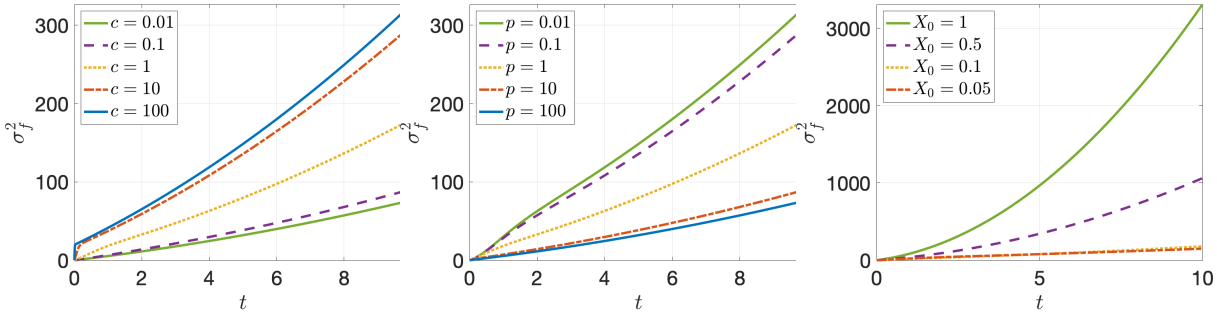
(a) Varying c (b) Varying p (c) Varying X_0

Figure 4.3: Monomer concentrations over time, varying the filament branching rate c, filament extension rate p and initial NLR concentration X_0. Results from numerical solutions of ODE model are shown. Initial ASC monomer concentration $\hat{x}_0 = 1$ is fixed, and $c = p = 1$ and $X_0 = 0.1$ unless stated otherwise. Monomer concentration is depleted in an exponential fashion, approaching 0 at large t. The rate of depletion is faster for higher values of c and p, and for higher initial numbers of NLR oligomers.



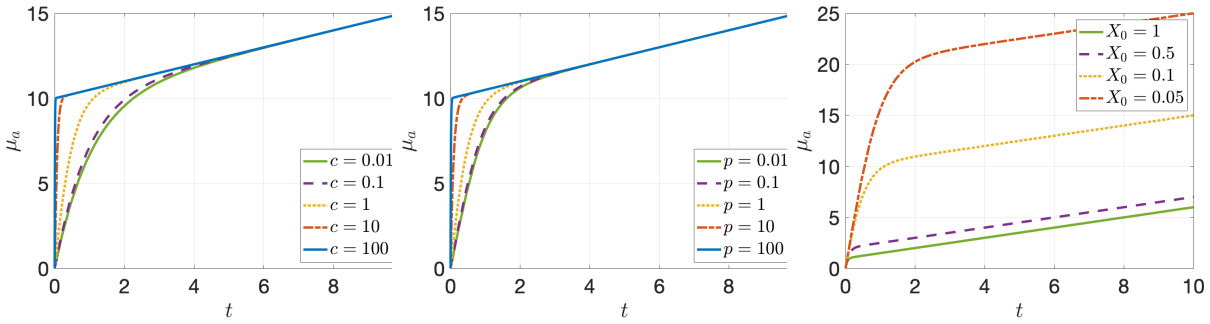
(a) Varying c (b) Varying p (c) Varying X_0

Figure 4.4: Mean number of filaments per cluster over time, varying the filament branching rate c, filament extension rate p and initial NLR concentration X_0. Results from numerical solutions of ODE model are shown. Initial ASC monomer concentration $\hat{x}_0 = 1$ is fixed, and $c = p = 1$ and $X_0 = 0.1$ unless stated otherwise. Following an initial sharp increase, μ_f tends towards a straight line with positive gradient; the gradient increases with the value of X_0, and the intercept increases with increasing c and decreasing p. However, this suggests that the mean cluster size grows indefinitely, which is biologically implausible.



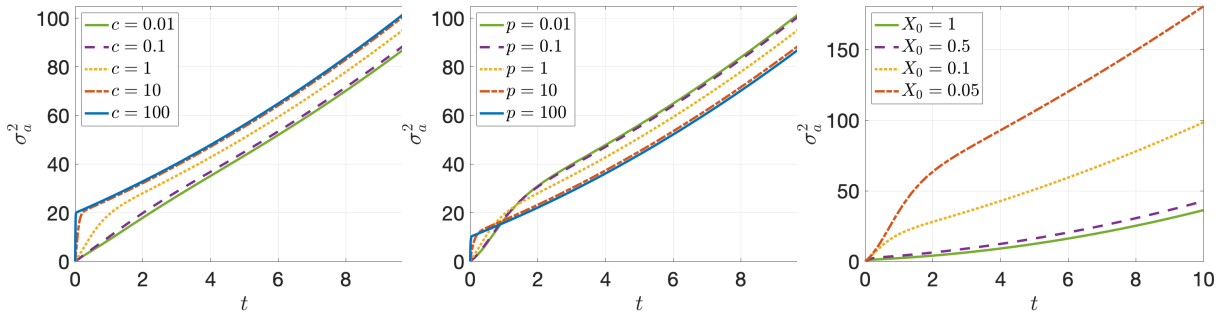
(a) Varying c (b) Varying p (c) Varying X_0

Figure 4.5: Variance of filaments per cluster over time, varying the filament branching rate c, filament extension rate p and initial NLR concentration X_0. Results from numerical solutions of ODE model are shown. Initial ASC monomer concentration $\hat{x}_0 = 1$ is fixed, and $c = p = 1$ and $X_0 = 0.1$ unless stated otherwise. At large times, the variance of the number of filaments in each cluster increases quadratically, when we would expect this to tend towards zero as a single cluster emerges.



(a) Varying c (b) Varying p (c) Varying X_0

Figure 4.6: Mean ASC monomers per cluster over time, varying the filament branching rate c, filament extension rate p and initial NLR concentration X_0. Results from numerical solutions of ODE model are shown. Initial ASC monomer concentration $\hat{x}_0 = 1$ is fixed, and $c = p = 1$ and $X_0 = 0.1$ unless stated otherwise. Following an initial sharp increase, μ_A tends towards a straight line with positive gradient; the gradient increases with decreasing X_0, suggesting the final growth rate is larger when there are fewer NLR oligomers initially present. However, once again this behaviour is biologically implausible as we would expect the mean cluster size to approach a constant value as monomers are depleted.

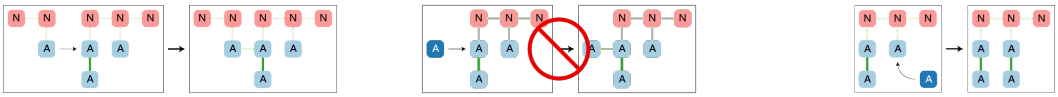


(a) Varying c (b) Varying p (c) Varying X_0

Figure 4.7: Variance of ASC monomers per cluster over time,varying the filament branching rate c, filament extension rate p and initial NLR concentration X_0. Results from numerical solutions of ODE model are shown. Initial ASC monomer concentration $\hat{x}_0 = 1$ is fixed, and $c = p = 1$ and $X_0 = 0.1$ unless stated otherwise. After an initial rapid increase, at large times, σ_A^2 increases linearly, when we would expect $\sigma_A^2 \to 0$ as monomers are depleted.

Variable	$t \ll 1$	$t \gg 1$
$\hat{x}$	$\hat{x}_0(1-(c+p)N_0X_0t)+O(t^2)$	$\frac{N\hat{x}_0e^{-Nt}}{(c+p)N_0X_0}+O(e^{-2Nt})$
X	$X_0\left(1-\frac{c_2X_0}{2}t\right)+O(t^2)$	$\frac{2}{c_2t}+O(t^{-2})$
μ_F	$N_0\left(1+\left(c\hat{x}_0+\frac{c_2X_0}{2}\right)t\right)+O(t^2)$	$\frac{N}{2X_0(c+p)}\left(c_2X_0t+2\right)+O(e^{-Nt})$
σ_F^2	$N_0\left(c\hat{x}_0+\frac{c_2N_0X_0}{2}\right)t+O(t^2)$	$\frac{N}{(c+p)^2N_0X_0^2}\left(\frac{c_2^2N_0X_0^2N}{2}t^2+c_2X_0\left(NN_0+c\hat{x}_0\right)t+2c\hat{x}_0\right)+O(te^{-Nt})$
μ_A	$N_0\hat{x}_0(c+p)t+O(t^2)$	$\frac{\hat{x}_0}{2cX_0}\left(c_2X_0t+2\right)+O(e^{-Nt})$
σ_A^2	$N_0\hat{x}_0(c+p)t+O(t^2)$	$\frac{\hat{x}_0}{N_0X_0}+\frac{(c+2p)\hat{x_0}}{c}+\frac{2pN}{c^2}\log\left(\frac{(c+p)N_0X_0}{N}\right)+\frac{c_2\hat{x}_0^2}{2}t+O(e^{-Nt})$

Table 4.1: Limits of solutions for the moments for large and small t, assuming all other parameters are $O(1)$.



(a) Reaction C2 included (b) Reaction C not included (c) Reaction P included

Figure 4.8: The simplified inflammasome model with filament extension only

to assess the behaviour of the model at early and late times, as for the model with branching and filament extension; in particular, we will be able to judge if either of these reactions is solely responsible for the unlimited growth of clusters observed in the previous section.

Models without branching

We will first consider the behaviour of the model with no branching, i.e. growth occurs by filament extension only (Figure 4.8).

The solutions for $\hat{x}, X, F$ and the generating function G are as follows:

$$\hat{x} = \hat{x}_0 e^{-pX_0N_0t} \tag{4.13a}$$

$$X = \frac{2X_0}{c_2X_0t + 2} \tag{4.13b}$$

$$F = X_0N_0 \tag{4.13c}$$

$$G = \frac{4X_0}{(c_2X_0t+2)^2 y^{-N_0} \exp[\hat{x}_0(z-1)(e^{-pX_0N_0t}-1)] - c_2X_0t(c_2X_0t+2)} \tag{4.13d}$$

The moments are as follows:

$$\mu_A = \frac{\hat{x}_0(1-e^{-N_0X_0pt})(X_0c_2t+2)}{2} \tag{4.14a}$$

$$\sigma_A^2 = \frac{\hat{x}_0 e^{-2N_0X_0pt}(e^{N_0X_0pt}-1)(X_0c_2t+2)((X_0c_2t\hat{x}_0+2)e^{N_0X_0pt} - X_0c_2t\hat{x}_0)}{4} \tag{4.14b}$$

$$\mu_F = \frac{N_0(X_0c_2t+2)}{2} \tag{4.14c}$$

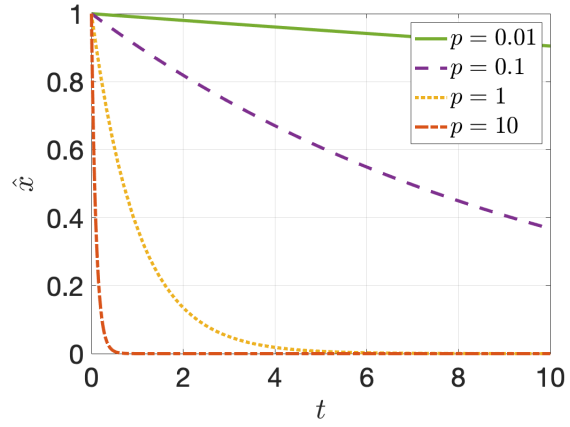
$$\sigma_F^2 = \frac{N_0^2X_0^2c_2t(X_0c_2t+2)}{4} \tag{4.14d}$$

The moments at early and late times, assuming other parameters are $O(1)$, are given in Table 4.2. Examples of the solutions for $\hat{x}, \mu_A$ and σ_A^2 are shown in Figure 4.9.

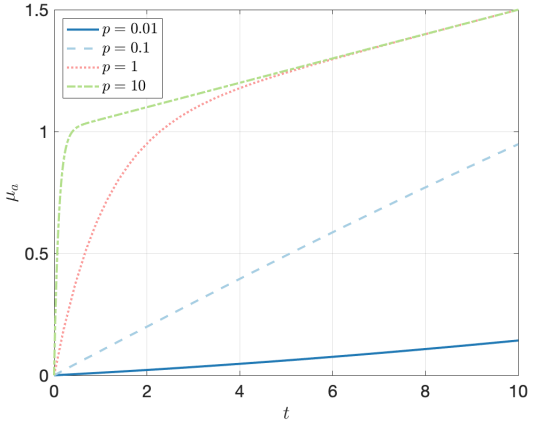
Model with branching only

We will first consider the behaviour of the model with branching only (Figure 4.10).

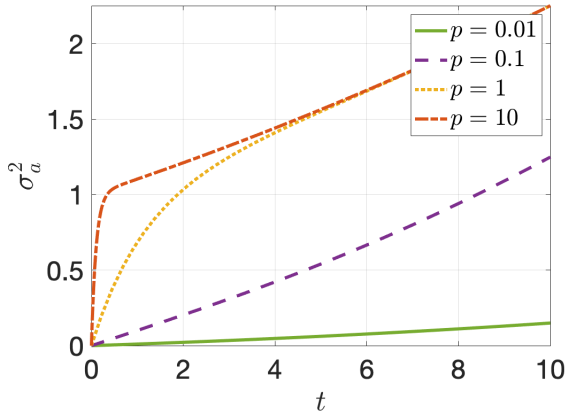
The solutions for $\hat{x}, X, F$ and the generating function G are as follows:



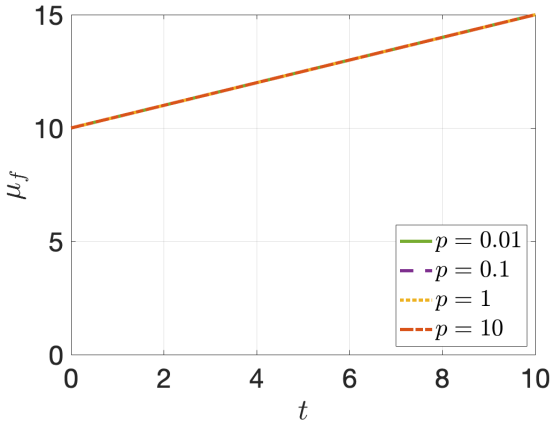
(a) ASC monomer concentration



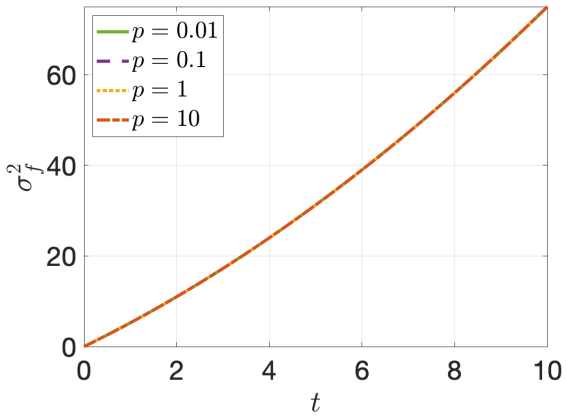
(b) μ_A



(c) σ_A^2



(d) μ_F

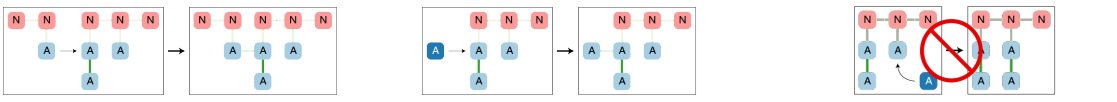


(e) σ_F^2

Figure 4.9: Moments for simplified model with no branching, varying the value of the filament extension rate p. Results from numerical solutions of ODE model are shown. Note that the mean and variance of the number of filaments per cluster (μ_F and σ_F^2) are independent of p, increasing linearly and quadratically respectively. Solutions are broadly similar to the solutions for the model with branching, but reactions generally occur more slowly, due to the lack of the positive feedback loop of filament creation caused by branching reactions. At late times, the variances of the number of filaments per cluster and the number of ASC monomers per cluster fail to tend to zero, so the model does not accurately capture inflammasome formation at late times.

Variable	$t \ll 1$	$t \gg 1$
$\hat{x}$	$\hat{x}_0(1 - pX_0N_0t) + O(t^2)$	$\hat{x}_0 e^{-pX_0N_0t}$
μ_F	$\frac{N_0(X_0c_2t+2)}{2}$	$\frac{N_0(X_0c_2t+2)}{2}$
σ_F^2	$\frac{N_0^2X_0^2c_2t}{2} + O(t^2)$	$\frac{N_0^2X_0^2c_2t(X_0c_2t+2)}{4}$
μ_A	$\hat{x}_0N_0X_0pt + O(t^2)$	$\frac{\hat{x}_0(X_0c_2t+2)}{2} + O(e^{-N_0X_0pt})$
σ_A^2	$\hat{x}_0N_0X_0pt + O(t^2)$	$\frac{\hat{x}_0(X_0c_2t+2)(\hat{x}_0X_0c_2t+2)}{4} + O(t^2e^{-N_0X_0pt})$

Table 4.2: Limits of solutions to the model with no branching for the moments for large and small t, assuming all other parameters are $O(1)$.



(a) Reaction C2 included (b) Reaction C included (c) Reaction P not included

Figure 4.10: The simplified inflammasome model

$$\hat{x} = \frac{\hat{x}_0(\hat{x}_0 + X_0N_0)}{\hat{x}_0 + X_0N_0e^{c(\hat{x}_0+X_0N_0t)}} \tag{4.15a}$$

$$X = \frac{2X_0}{c_2X_0t + 2} \tag{4.15b}$$

$$F = \frac{(\hat{x}_0 + X_0N_0)X_0N_0}{\hat{x}_0e^{-c(\hat{x}_0+X_0N_0)t} + X_0N_0} \tag{4.15c}$$

$$G = \frac{4X_0}{(c_2X_0t + 2)^2g^{-N_0} - c_2X_0t(c_2X_0t + 2)} \tag{4.15d}$$

$$g = \frac{\hat{x}_0 + X_0N_0}{\hat{x}_0e^{-c(\hat{x}_0+X_0N_0)t} + X_0N_0}\left(\frac{1}{y} - z\right) + z \tag{4.15e}$$

The moments are as follows:

$$\mu_A = \frac{N_0\hat{x}_0(e^{ct(\hat{x}_0+X_0N_0)} - 1)(X_0c_2t + 2)}{2(\hat{x}_0 + N_0X_0e^{ct(\hat{x}_0+N_0X_0)})} \tag{4.16a}$$

$$\sigma_A^2 = \frac{N_0\hat{x}_0(e^{ct(\hat{x}_0+N_0X_0)} - 1)(X_0c_2t + 2)((2\hat{x}_0 + 2N_0X_0 + N_0X_0c_2t\hat{x}_0)e^{ct(\hat{x}_0+N_0X_0)} - N_0X_0c_2t\hat{x}_0)}{4(\hat{x}_0 + N_0X_0e^{ct(\hat{x}_0+N_0X_0)})^2} \tag{4.16b}$$

$$\mu_F = \frac{N_0e^{ct(\hat{x}_0+N_0X_0)}(\hat{x}_0 + N_0X_0)(X_0c_2t + 2)}{2(\hat{x}_0 + N_0X_0e^{ct(\hat{x}_0+N_0X_0)})} \tag{4.16c}$$

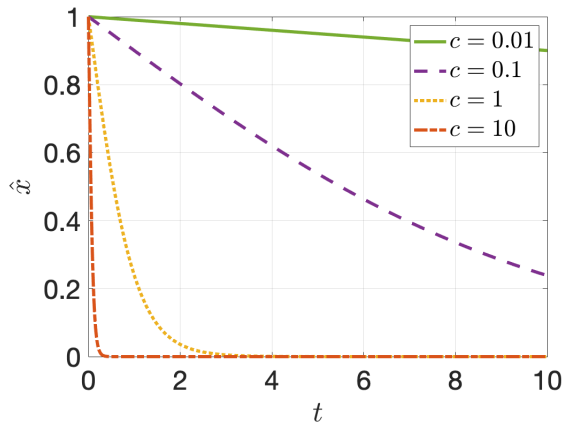
$$\sigma_F^2 = \frac{N_0e^{ct(\hat{x}_0+N_0X_0)}(\hat{x}_0 + N_0X_0)(X_0c_2t + 2)((2\hat{x}_0 + N_0^2X_0^2c_2t + N_0X_0c_2t\hat{x}_0)e^{ct(\hat{x}_0+N_0X_0)} - 2\hat{x}_0)}{4(\hat{x}_0 + N_0X_0e^{ct(\hat{x}_0+N_0X_0)})^2} \tag{4.16d}$$

The moments at early and late times, assuming other parameters are $O(1)$, are given in Table 4.3.

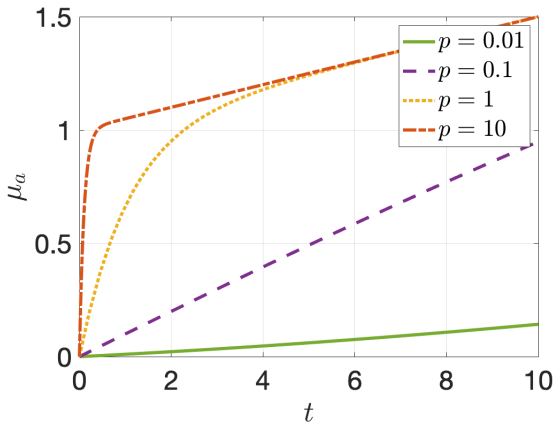
Comparing the models without branching and with branching only

The solutions for the models without branching and with branching only are broadly similar to the solutions for the model with both branching and filament extension; at small times, we see an approximately linear decrease in $\hat{x}$, and a linear increase in μ_F, σ_F^2, μ_A and σ_A^2. X is unchanged, since the cluster joining process is independent of the filament branching and extension reactions. At later times, as for the model with both branching and filament extension, we see that μ_F and μ_A increase approximately linearly, while σ_F^2 and σ_A^2 increase approximately quadratically. Once again, the fact that the mean does not tend to a fixed value, and the variance does not tend to zero, flout the conditions on succesful modelling of the inflammasome at late times.

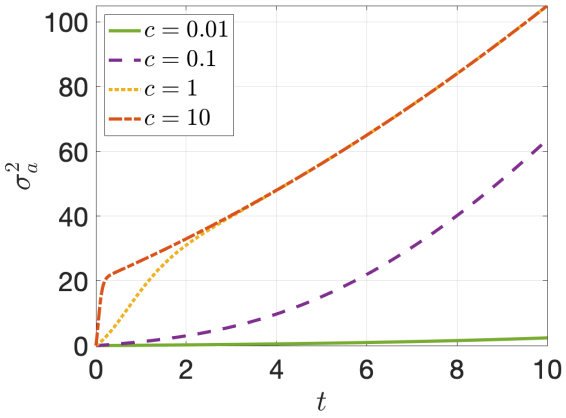
Interestingly, the solutions for μ_F and σ_F^2 in the model without branching are independent of the filament extension rate p; the total number of filaments in this model are constant, so



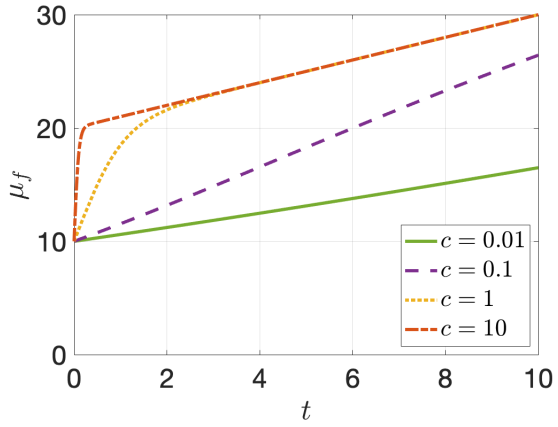
(a) ASC monomer concentration



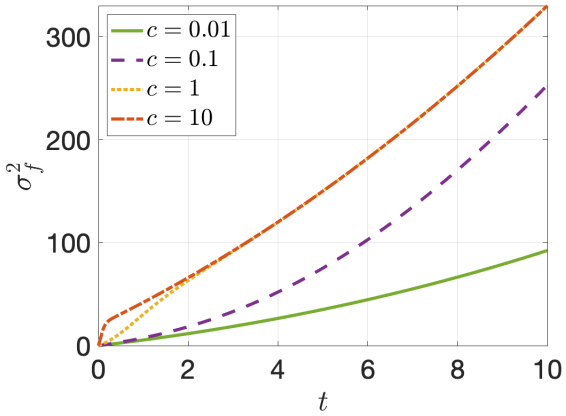
(b) μ_A



(c) σ_A^2



(d) μ_F



(e) σ_F^2

Figure 4.11: Moments for simplified model with branching growth only, varying the value of the filament branching rate c. Results from numerical solutions of ODE model are shown. Solutions are broadly similar to the full model, with initial rapid linear increases in cluster size for large c. Once again, the variances of the number of filaments per cluster and the number of ASC monomers per cluster fail to tend to zero, so the model fails at late times.

Variable	$t \ll 1$	$t \gg 1$
$\hat{x}$	$\hat{x}_0(1 - cX_0N_0t) + O(t^2)$	$\frac{(\hat{x}_0+N_0X_0)\hat{x}_0e^{-ct(\hat{x}_0+N_0X_0)}}{N_0X_0} + O(e^{-2ct(\hat{x}_0+N_0X_0)})$
μ_F	$\frac{N_0(X_0c_2t+2c\hat{x}_0t+2)}{2}$	$\frac{\hat{x}_0(X_0c_2t+2)}{2X_0} + O(te^{-ct(\hat{x}_0+N_0X_0)})$
σ_F^2	$\frac{N_0(N_0X_0c_t+2c\hat{x}_0)t}{2} + O(t^2)$	$\frac{\hat{x}_0(X_0c_2t+2)(2\hat{x}_0+2N_0X_0+N_0X_0c_2t\hat{x}_0)}{4N_0X_0^2}$ + $O(t^2e^{-ct(\hat{x}_0+N_0X_0)})$
μ_A	$\hat{x}_0N_0ct + O(t^2)$	$\frac{(\hat{x}_0+N_0X_0)(X_0c_2t+2)}{2X_0} + O(t^2e^{-ct(\hat{x}_0+N_0X_0)})$
σ_A^2	$\hat{x}_0N_0ct + O(t^2)$	$\frac{(\hat{x}_0+N_0X_0)(X_0c_2t+2)(2\hat{x}_0+N_0^2X_0^2c_2t+N_0X_0c_2\hat{x}_0t)}{4N_0X_0^2}$ + $O(t^2e^{-ct(\hat{x}_0+N_0X_0)})$

Table 4.3: Limits of solutions to the model with branching growth only for the moments for large and small t, assuming all other parameters are $O(1)$.

the distribution of filaments per cluster is only affected by the total number of clusters. As the total number of clusters decreases, the mean number of filaments per cluster increases linearly, and the variance increases quadratically.

More specifically, at small times, we see that $\hat{x}$ is broadly unaffected by switching between branching and non-branching behaviours (if p and c to take the same values). However, the initial rate of increase of μ_F and σ_F^2 is slower in the case of the model without branching compared to the model with branching; in the model with branching, the initial rate of linear increase has an additional factor of $\hat{x}_0N_0c$. For small initial concentrations X_0 we also see a much slower initial increase in μ_A and σ_A^2 for the model without branching. Once again, these differences in rates arise from the positive feedback loop of filament creation, which arises in the model with branching but not in the model without branching.

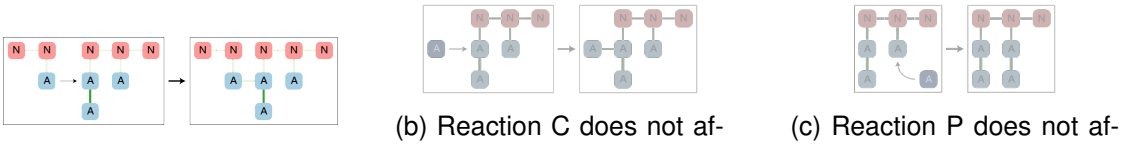
At later times, the faster increase in cluster growth rate for the branching-only model is also reflected in the rate of monomer deletion, with an extra factor of $-c\hat{x}_0t$ appearing in the exponent. In general, the values tended to for μ_F, μ_A, σ_F^2 and σ_A^2 in the model with branching are also much higher than in the model without branching; once again, this reflects the

postive feedback effect when branching occurs.

4.3 Analysis of the stochastic model

So far, I have focused on using deterministic methods to analyse the simplified inflammasome model; this has given an insight into the roles of the cluster growth and joining processes, and in particular, how the number of free ASC monomers and distribution of cluster sizes changes over time for different parameter values. However, it is evident that the deterministic models do not capture some aspects of inflammasome formation which would be useful to investigate further. For example, the deterministic model fails when the number of reactants is very small, i.e. at late stages in the process. The deterministic model also gives us no insight into variation in the inflammasome formation process between cells in the same population. This motivates a discrete stochastic approach, which addresses both of these issues. A stochastic system which tracks the size of all clusters over time would be extremely unwieldy; the relative simplicity of analysing the deterministic system makes it a more appropriate approach to considering distributions of cluster size. However, it is still possible to analyse stochastic systems involving the 'sum variables' introduced in Chapter 3; namely, $F(t)$, $\hat{x}(t)$ and $X(t)$ (the total number of filaments, free ASC monomers and clusters at time t respectively); note that since we are now working with a discrete stochastic system, these variables now correspond to total abundances of reactants rather than concentrations. In particular, stochastic analysis of $X(t)$ can allow us to determine the distribution of the time at which the first joining event occurs between clusters, and the time at which a single isolated cluster first emerges; neither of these can be obtained from the deterministic model, but distributions of both times can be analytically obtained from the stochastic model using a CTMC approach. Analysing $\hat{x}(t)$ is also useful since this is the quantity that can most easily be compared to the experimental data presented in Chapter 2; in particular, I will present analytical expressions for the probability density functions of the relevant characteristic times.

One particular benefit of using the simplified inflammasome model for stochastic analysis is that the cluster joining process is independent of cluster growth; therefore, these two processes can be considered separately. Allowing cluster joining rates to be dependent on cluster size adds considerably to the complexity of the stochastic model; this will be investigated further using simulations in Chapter 5.



(a) Reaction C2 included

(b) Reaction C does not affect this subsystem

(c) Reaction P does not affect this subsystem

Figure 4.12: The cluster joining subsystem for the simplified inflammasome model

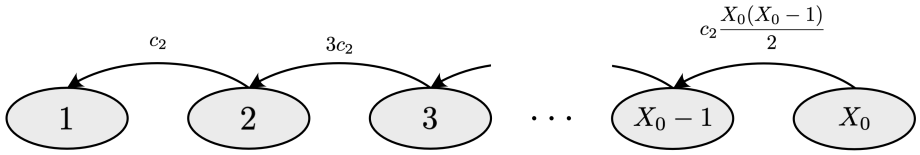


Figure 4.13: The number of clusters $X(t)$ forms a continuous-time Markov chain, with transition rates proportional to the number of joins possible with the clusters present, and independent of cluster size.

4.3.1 Stochastic model for number of clusters

We will first consider the subsystem of the model with cluster joining reactions only (Figure 4.12).

In the simplified model, the rate of cluster joining does not depend on cluster size; the rate at which joining events occur is proportional to the number ways a join could occur between two clusters. When there are n clusters, the rate of joining is therefore proportional to $\binom{n}{2} = \frac{n(n-1)}{2}$.

The number of clusters $X(t)$ can be considered as a CTMC, shown in Figure 4.13. Note that the single cluster state is absorbing.

We can calculate the distribution of the first joining and single cluster emergence times as first hitting times for states $X_0 - 1$ and 1 in the CTMC. The time taken to jump from one state $X = n$ to $X = n-1$ in our Markov chain is exponentially distributed with rate $e^{c_2 n(n-1)/2}$. For example, the first joining time is exponentially distributed with rate $c_2\frac{X_0(X_0-1)}{2}$. The mean first cluster joining time is $\frac{2}{c_2 X_0(X_0-1)}$, and the variance is $\frac{4}{(c_2 X_0(X_0-1))^2}$.

The time to single cluster emergence is the sum of the initial jump times from $X = X_0$ to $X = X_0 - 1$, $X = X_0 - 1$ to $X = X_0 - 2, \cdots, X = 2$ to $X = 1$, each of which is exponentially distributed; the total time to cluster emergence is therefore the sum of exponentially distributed random variables with rates $c_2, 3c_2, \cdots, c_2 X_0(X_0-1)/2$. By definition, this random variable is hypoexponentially distributed with parameters $c_2, 3c_2, \cdots, c_2 X_0(X_0-1)/2$ [Amari

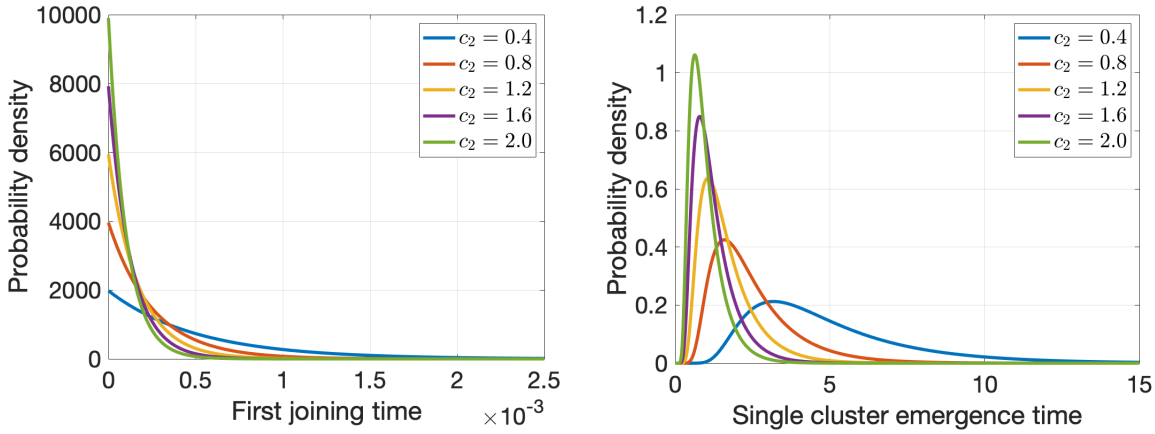


Figure 4.14: Analytically-derived distributions of times of first joining event and single cluster emergence while varying the cluster joining rate c_2, fixing the initial NLR oligomer concentration $X_0 = 100$. Note that units of time are scaled by the rate c_2. Increasing c_2 decreases both the mean and the spread of both distributions; for all values of c_2 the distribution has a long tail corresponding to rare events at late times.

and Misra, 1997].

A hypoexponential distribution with parameters $\lambda_1, \lambda_2, \cdots, \lambda_n$ has mean $\sum_{i=1}^{n} 1/\lambda_i$, and variance $\sum_{i=1}^{n} 1/\lambda_i^2$. Provided the λ_i are unique, the probability density function is as follows:

$$f(x) = \sum_{i=1}^{n} \lambda_i e^{-\lambda_i x} \prod_{j=1, j\neq i}^{n} \frac{\lambda_j}{\lambda_j - \lambda_i} \tag{4.17}$$

The mean time of single cluster emergence is therefore $\frac{2}{c_2}\left(1 - \frac{1}{X_0}\right)$, and the variance is $\frac{2}{c_2}\sum_{n=1}^{X_0-1} \frac{1}{n^2(n+1)^2} \approx \frac{2}{3c_2}(\pi^2 - 9)$ for large X_0. The probability density function $f_{T_s}(t)$ is as follows:

$$\begin{aligned} f_{T_s}(t) &= c_2 \sum_{i=1}^{X_0-1} \frac{i(i+1)}{2} e^{-c_2 i(i+1)t/2} \left(\prod_{j=1, j\neq i}^{X_0-1} \frac{c_2 j(j+1)/2}{c_2 j(j+1)/2 - c_2 i(i+1)/2} \right) \\ &= \frac{c_2}{2}(X_0 - 1)!\, X_0! \sum_{i=1}^{X_0-1} e^{-c_2 i(i+1)t/2} \frac{(-1)^{i-1} i(i+1)(2i+1)}{(X_0 - i - 1)!\,(X_0 + i)!} \end{aligned} \tag{4.18}$$

The only parameters which affect first joining time and single cluster emergence time in this model are c_2 and X_0 (Figures 4.14 and 4.15). The distributions for both times also show considerable positive skew, so all have a long tail corresponding to the increasing rarity of events at later times. As c_2 increases, the distributions of first joining and single cluster emergence times decrease both in mean and spread. Hence, for fast joining rates, unsurprisingly cluster formation is faster, and formation of clusters at later times is extremely rare. Interestingly, varying X_0 has fairly little relative effect on the distribution of single cluster emergence

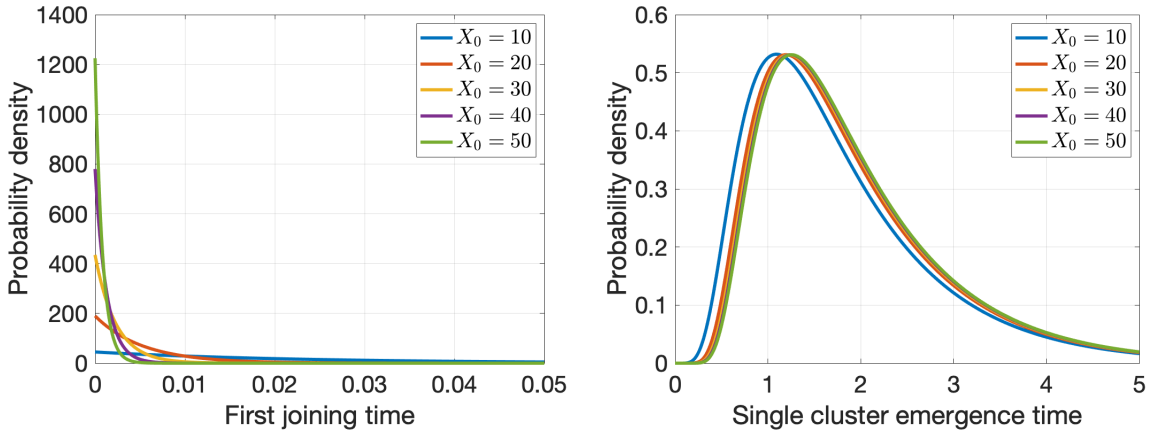


Figure 4.15: Analytically-derived distributions of times of first joining event and single cluster emergence while varying the initial NLR oligomer concentration X_0, fixing the cluster joining rate $c_2 = 1$. Note that units of time are scaled by the rate c_2. Increasing X_0 decreases both the mean and the spread of the distribution of first joining times; however, varying X_0 has little effect on the time of single cluster emergence.

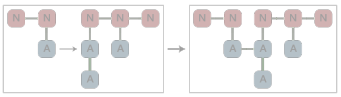
times, although increasing X_0 leads to a slight increase in single cluster emergence times; the mean single cluster emergence time slowly approaches a limit of $2/c_2$ as X_0 increases. However, first joining time is extremely sensitive to the value of X_0; larger values of X_0 lead to much faster first joining times, with reduced distribution spread.

This model therefore shows that although the first joining time is sensitive to the initial number of NLR oligomers, this parameter does not have a notable effect on the distribution of inflammasome formation time. This suggests the variation between cells forming inflammasomes in similar conditions arises from other sources, possibly due to a difference in cluster joining rate. However, this still bears further investigation; in this simplified model, I have decoupled the cluster joining and cluster growth processes, but this conclusion cannot be extended to models with size-dependent cluster joining. Since such joining processes are considerably more involved to analyse analytically, I will examine them further using simulations in Chapter 5.

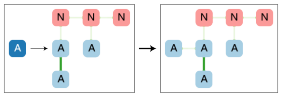
4.3.2 Monomer depletion

We now consider the subsystem of the simplified inflammasome model corresponding to monomer depletion reactions (Figure 4.16).

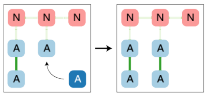
Let $F(t)$ and $\hat{X}(t)$ be the total number of ASC filaments in the system and the number of unbound ASC monomers at time t respectively. We can then view $F(t)$ and $\hat{X}(t)$ as a CTMC starting from initial conditions $F(0) = f_0$ and $\hat{X}(0) = \hat{x}_0$.



(a) Reaction C2 does not affect this subsystem



(b) Reaction C included



(c) Reaction P included

Figure 4.16: The cluster joining subsystem for the simplified inflammasome model

Reference	Reaction	Propensity function	Explanation
C	$(\hat{X}, F) \xrightarrow{c} (\hat{X}-1, F+1)$	$c\hat{X}F$	ASC monomer in cytosolic compartment crosslinks to filament, forms new filament
P	$(\hat{X}, F) \xrightarrow{p} (\hat{X}-1)$	$p\hat{X}F$	ASC monomer in cytosolic compartment extends existing filament

Table 4.4: Reactions considered in the stochastic filament extension/branching and monomer depletion subsystem.

There are two reactions to consider, set out in Table 4.4.

We can regard the process of monomer depletion as a series of jumps on a lattice in which each point corresponds to a value of $\hat{X}$ and F (Figure 4.17); the times between jumps are all exponentially distributed

The two possible jumps out of a point $(\hat{X}, F)$ are a jump caused by filament extension to point $(\hat{X}-1, F)$ with rate $pF\hat{X}$, and a jump caused by filament branching to point $(\hat{X}-1, F+1)$ with rate $cF\hat{X}$.

At any point, the next jump will cause filament branching with probability $\frac{c\hat{X}F}{c\hat{X}F+p\hat{X}F} = \frac{c}{c+p}$, independent of the current number of filaments or monomers in the system, the time of the jump, and of any previous jumps. Thus, the total number of filaments once all monomers have been depleted will be binomially distributed with $\hat{x}_0$ total trials and success probability $\frac{c}{c+p}$.

Furthermore, we can investigate the monomer depletion characteristic times $t_{1/q}$ for the monomer population to be depleted by a proportion $1/q$. Note that the time taken for N monomers to be depleted (or N jumps to be taken through the lattice in Figure 4.17) is $t_{N/\hat{x}_0}$.

We can consider the distribution of $t_{N/\hat{x}_0}$ as a phase-type distribution [Neuts, 1975]; this

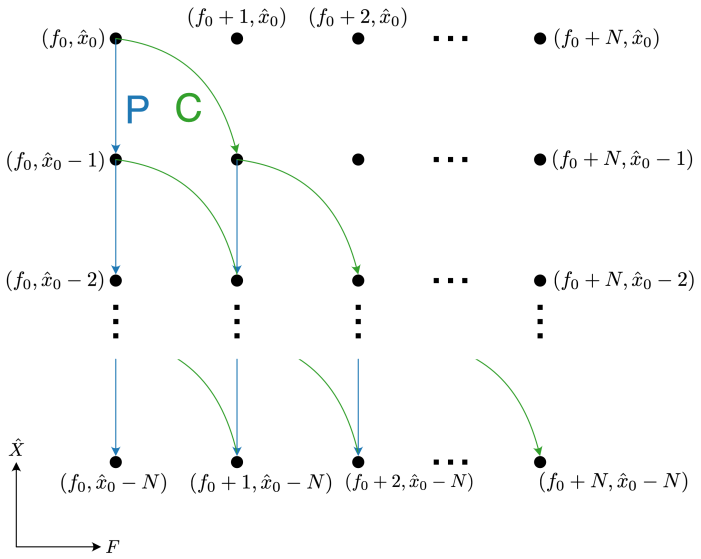


Figure 4.17: The process of monomer depletion and filament formation and extension can be represented as movement between points on a lattice. Here, the row represents the number of free ASC monomers and the column represents the total number of filaments. There are two types of jumps between lattice points - jumps representing the P reaction, in which the number of monomers $\hat{X}$ decreases by 1, which occur at the rate $pF\hat{X}$, and jumps representing the C reaction, in which $\hat{X}$ decreases by 1 and F increases by 1, which occur at the rate $cF\hat{X}$.

is defined as the time until absorption of a CTMC with an absorbing state. The distribution is defined by a $n \times n$ generator matrix S of transition rates between the non-absorbing states in the CTMC, and a vector π of probabilities of the system starting in each state. The full $(n+1) \times (n+1)$ CTMC then has the following generating matrix:

$$\begin{pmatrix} S & \mathbf{s} \\ \mathbf{0} & 0 \end{pmatrix} \tag{4.19}$$

where $\mathbf{s}$ is a $n \times 1$ vector of absorption rates and $\mathbf{0}$ is a $1 \times n$ vector of zeros. Note that the hypoexponential distribution is a special case of a phase-type distribution, with S taking the following form:

$$\begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & -\lambda_2 & \lambda_2 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & -\lambda_{n-2} & \lambda_{n-2} & 0 \\ 0 & 0 & \cdots & 0 & -\lambda_{n-1} & \lambda_{n-1} \\ 0 & 0 & \cdots & 0 & 0 & -\lambda_n \end{pmatrix} \tag{4.20}$$

where λ_i is the ith jump rate.

In the case of the monomer depletion and filament growth/branching system , we will define an ordering of states as follows: $(f_0, \hat{x}_0), (f_0+1, \hat{x}_0), (f_0+2, \hat{x}_0), \cdots, (f_0+N, \hat{x}_0), (f_0, \hat{x}_0 - 1), (f_0+1, \hat{x}_0-1), \cdots, (f_0+N, \hat{x}_0-N)$. The ith state then has $\hat{X}(i) = \left\lfloor \frac{i}{N+1} \right\rfloor$ and $F(i) = f_0 + (i-1) \mod (N+1)$

We group all states with $\hat{X} \leq \hat{x}_0 - N$ into a single absorbing state. Then there are $N(N+1)$ non-absorbing states, and the generating matrix is

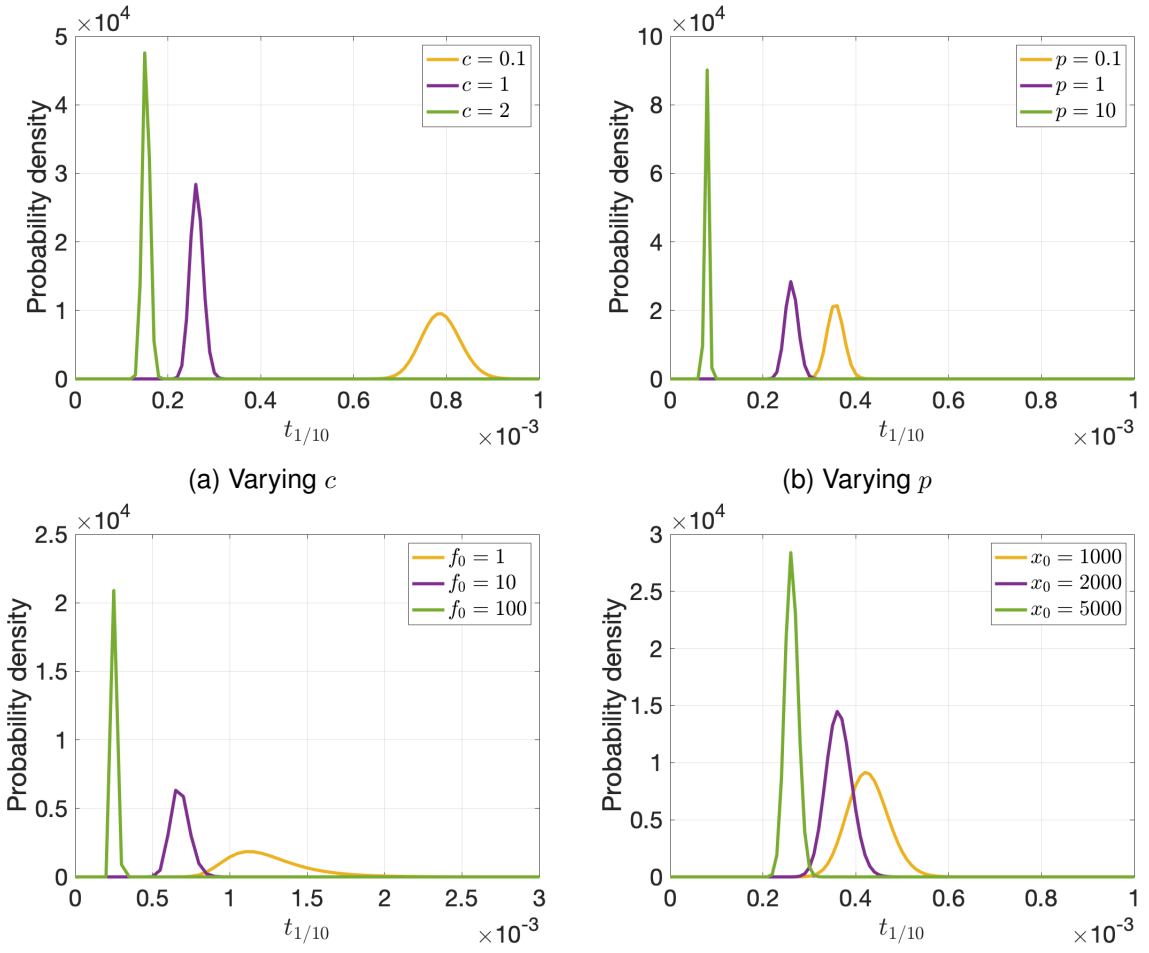
$$S = \begin{pmatrix} -(c+p)d_1 & 0 & \cdots & pd_1 & cd_1 & 0 & \cdots & 0 \\ 0 & -(c+p)d_2 & \cdots & 0 & pd_2 & cd_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -(c+p)d_{N^2-2} & 0 & \cdots & pd_{N^2-2} & cd_{N^2-2} \\ 0 & 0 & \cdots & 0 & -(c+p)d_{N^2-1} & 0 & \cdots & pd_{N^2-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & -(c+p)d_{N^2(N+1)^2-1} & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & -(c+p)d_{N^2(N+1)^2} \end{pmatrix} \tag{4.21}$$

where $d_i = F(i)(\hat{x}_0 - \hat{X}(i) + 1)$ for state i. This matrix is upper triangular, and the only non-zero entries lie on the main diagonal and the $(N+1)$th and $(N+2)$th diagonals. Since we start in the state $(f_0, \hat{x}_0)$, we put $\pi = (1, 0, \cdots, 0)^T$.

Note that not all states can be reached from this starting condition; in particular, states with $F - f_0 > \hat{x} - \hat{X}$ cannot be reached since these would correspond to states in which more filaments have been created than monomers depleted.

The random variable $t_{N/\hat{x}_0}$ has probability density function $f_{t_{N/\hat{x}_0}}(t) = \pi e^{tS}\mathbf{s}$ [Neuts, 1975]; some examples for different parametrisations are given in Figure 4.18. As far as I have been able to determine, this cannot be expressed in a simplified closed form but can be estimated using computational techniques; in this case, I have used the Matlab `fastexp` package to calculate the pdf for a sparse representation of Matrix 4.21 [Higham, 2021].

We can also calculate the mean and variance of this phase-type distribution; the mean of $t_{N/\hat{x}_0}$ is given by $-\pi S^{-1}\mathbf{1}$, where $\mathbf{1}$ is a vector of ones, and the variance is given by $2\pi S^{-2}\mathbf{1} - (\pi S^{-1}\mathbf{1})^2$ [Neuts, 1975]. Some example values are given in Figure 4.19. The mean and variance of $t_{N_0/\hat{x}_0}$ decrease with increasing c, p, f_0 and $\hat{x}_0$. This is unsurprising in the case of c and p since this increases the overall rate at which ASC monomers are depleted; however, notably, increasing c leads to a much larger reduction in mean and variance than increasing p by the same amount. This is due to the positive feedback loop caused by C reactions creating more filaments, which further increases the rate of depletion of monomers. Increasing f_0



(a) Varying c

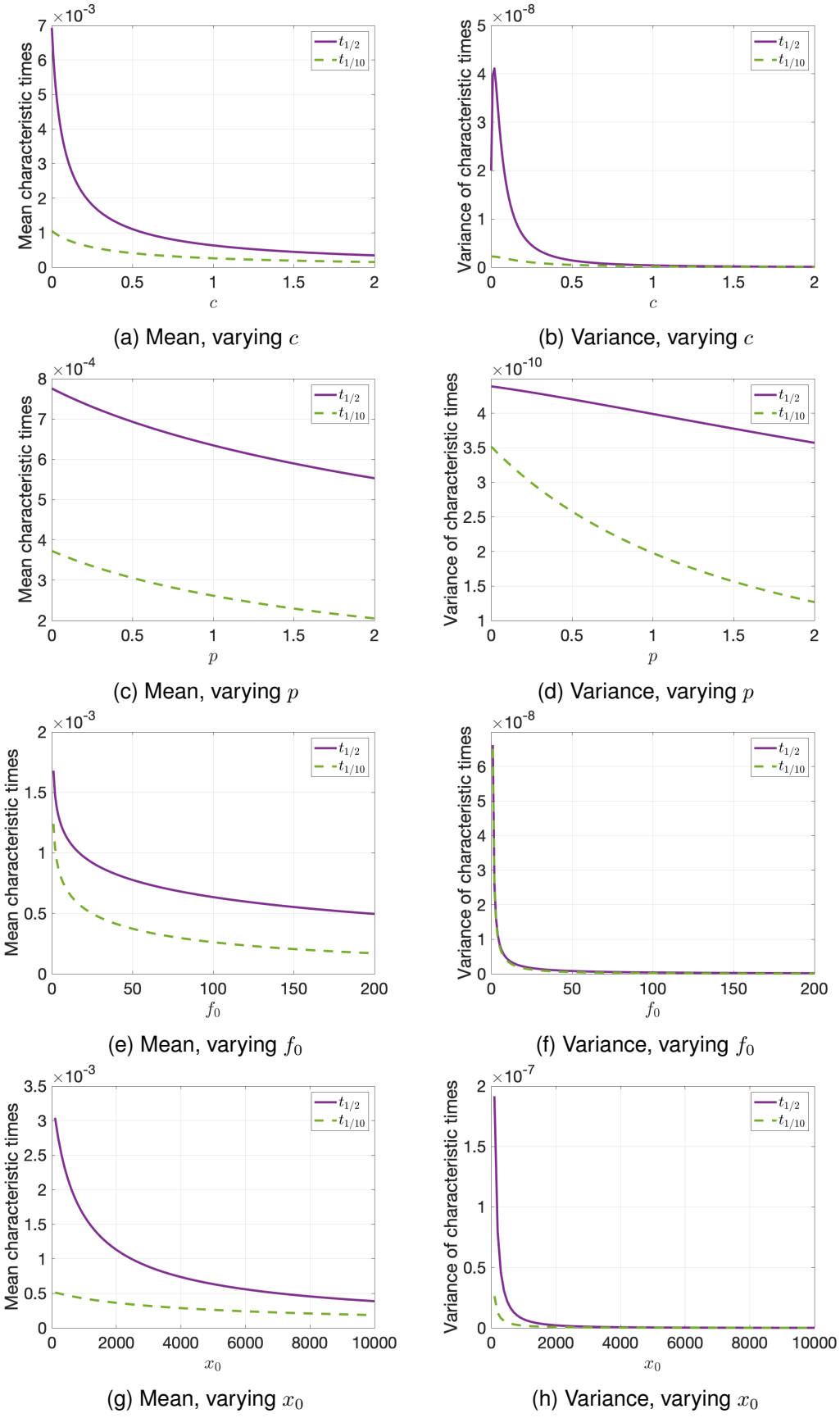
(b) Varying p

(c) Varying f_0

(d) Varying x_0

Figure 4.18: Analytically-derived Probability density functions of the monomer depletion characteristic time $t_{1/10}$ calculated for varying values of the branching rate c, filament extension rate p, and initial number of NLRs and ASC monomers, f_0 and x_0 respectively. When values are not given explicitly, $c = p = 1, f_0 = 100, x_0 = 5000$. As c, p, f_0 and x_0 increase, the mean and spread of the distributions of $t_{1/10}$ decrease.

and $\hat{x}_0$ decreases the mean and variance of the distribution of $t_{N/\hat{x}_0}$, since increasing these parameters increases the initial rate of ASC monomer depletion. Another notable feature of these distributions is that, unlike the single cluster emergence times, they display very little skewness; this may be due to the fact that the time for the monomer population to decrease by a given fraction is not as sensitive to individual reactions taking a longer time compared to single cluster emergence (which is particularly sensitive to the time taken for the last few reactions, when few clusters remain).



(a) Mean, varying c

(b) Variance, varying c

(c) Mean, varying p

(d) Variance, varying p

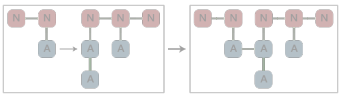
(e) Mean, varying f_0

(f) Variance, varying f_0

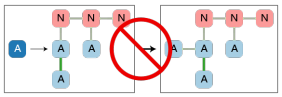
(g) Mean, varying x_0

(h) Variance, varying x_0

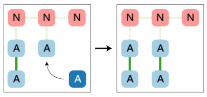
Figure 4.19: Analytically-derived means and variances of $t_{1/10}$ and $t_{1/2}$ while varying the filament branching rate c, filament extension rate p, initial number of filaments f_0 and initial number of ASC monomers x_0; other variables are fixed at $c = 1, p = 1, f_0 = 100, X_0 = 5000$. As c, p f_0 and x_0 increase, mean and variance decrease. This effect is especially marked for the branching rate c, since as c increases this reinforces the positive feedback loop of increasing rate of monomer depletion as more branches are formed.



(a) Reaction C2 does not affect this subsystem



(b) Reaction C is not included



(c) Reaction P is included

Figure 4.20: The simplified inflammasome model with filament extension only

Model without branching

The model becomes much simpler if branching (C) reactions are not included, as is currently the generally accepted conceptual model of inflammasome formation (Figure 4.20); while the time of single cluster emergence is not affected, since this is only dependent on the rate of cluster joining and the initial number of NLR oligomers, the distribution of characteristic times for monomer distribution $t_{N/\hat{x}_0}$ are affected. In fact, if N monomers are depleted, this will occur via a series of exponentially distributed jumps down the left-hand side of the lattice in Figure 4.17 with rates $pf_0\hat{x}_0, pf_0(\hat{x}_0-1), pf_0(\hat{x}_0-2), \cdots, pf_0\hat{x}_0-N+1)$. Thus $t_{1/n}$ will be hypoexponentially distributed, with rates $pf_0\hat{x}_0, pf_0(\hat{x}_0-1), \cdots, pf_0(\hat{x}_0-N+1)$.

The probability density function, mean and variance are as follows:

$$\mathbb{E}(t_{N/\hat{x}_0}) = \sum_{i=1}^{N} \frac{1}{pf_0(\hat{x}_0-i+1)} = \frac{1}{pf_0}(H_{\hat{x}_0} - H_{\hat{x}_0-N})) \tag{4.22a}$$

$$var(t_{N/\hat{x}_0}) = \sum_{i=1}^{N} \frac{1}{p^2f_0^2(\hat{x}_0-i+1)^2} = \frac{1}{p^2f_0^2}(H^{(2)}_{\hat{x}_0} - H^{(2)}_{\hat{x}_0-N}) \tag{4.22b}$$

$$\begin{aligned} f_{t_{N/\hat{x}_0}}(t) &= \sum_{i=1}^{N} pf_0(\hat{x}_0-i+1)e^{-pf_0(\hat{x}_0-i+1)t} \\ &\times \left(\prod_{j=1,j\neq i}^{N} \frac{pf_0(\hat{x}_0-j+1)}{pf_0(\hat{x}_0-j+1) - pf_0(\hat{x}_0-i+1)} \right) = pf_0N\binom{\hat{x}_0}{N} e^{-pf_0\hat{x}_0t}(e^{pf_0t} \\ &- 1)^{N-1} \end{aligned} \tag{4.22c}$$

where H_i is the ith harmonic number, $H_i = \sum_{j=0}^{i} \frac{1}{j}$ and $H_i^{(n)}$ is the ith generalised harmonic number $H_i^{(n)} = \sum_{j=0}^{i} \frac{1}{j^n}$. This expression is more amenable to fitting to experimental data than the full phase-type distribution. Example plots of the probability density function are given in Figure 4.21; the mean and variance of $t_{N/\hat{x}_0}$ decrease as p increases.

Model with branching but not filament extension

A similar analysis can be carried out in the case that no filament extension (P) reactions occur and inflammasome formation occurs through branching reactions only (Figure 4.22);

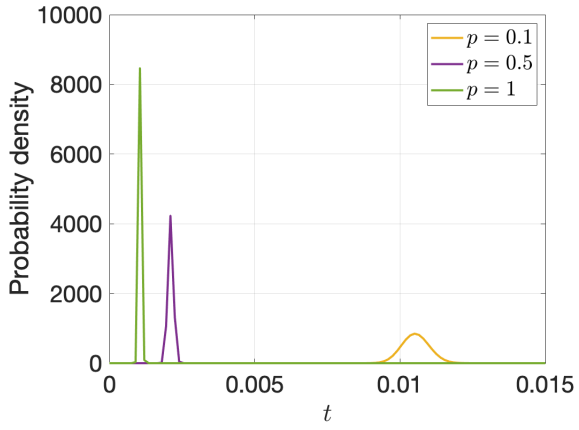
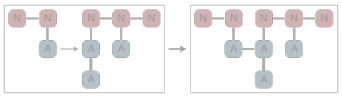
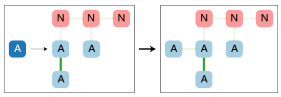


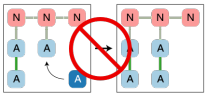
Figure 4.21: Analytically-derived examples of distributions for the characteristic monomer depletion time $t_{1/10}$, for the model without branching, for different values of p, the rate of branching. As p increases, $t_{1/10}$ decreases.



(a) Reaction C2 does not affect this subsystem



(b) Reaction C is included



(c) Reaction P is not included

Figure 4.22: The simplified inflammasome model with branching only

the system undergoes a series of diagonal jumps through the lattice in Figure 4.17. The times between jumps are exponentially distributed, with rates $cf_0\hat{x}_0, c(f_0+1)(\hat{x}_0-1), c(f_0+2)(\hat{x}_0-2), \cdots, c(f_0+N-1)(\hat{x}_0-N+1)$. Thus $t_{N/\hat{x}_0}$ will be hypoexponentially distributed, with rates $cf_0\hat{x}_0, c(f_0+1)(\hat{x}_0-1), \cdots, c(f_0+N-1)(\hat{x}_0-N+1)$.

The mean and variance are as follows:

$$\mathbb{E}(t_{N/\hat{x}_0}) = \sum_{i=1}^{N} \frac{1}{c(f_0+i-1)(\hat{x}_0-i+1} = \frac{1}{c(\hat{x}_0+f_0)}(H_{f_0+N-1} - H_{f_0-1} + H_{\hat{x}_0} - H_{\hat{x}_0-N})) \tag{4.23a}$$

$$\begin{aligned} var(t_N) &= \sum_{i=1}^{N} \frac{1}{c^2(f_0+i-1)^2(\hat{x}_0-i+1)^2} \\ &= \frac{1}{c^2(f_0+\hat{x}_0)^3}[2(H_{f_0+N-1} - H_{f_0-1} + H_{\hat{x}_0} - H_{\hat{x}_0-N}) \\ &\qquad + (f_0+\hat{x}_0)(H^{(2)}_{\hat{x}_0} - H^{(2)}_{\hat{x}_0-N} + H^{(2)}_{f_0+N-1} - H^{(2)}_{f_0-1}] \end{aligned} \tag{4.23b}$$

If $2N \geq \hat{x}_0 - f_0$, the rates are not unique; in particular, two rates $\lambda_i = c(f_0+i-1)(\hat{x}_0-i+1)$ and $\lambda_j = c(f_0+j-1)(\hat{x}_0-j+1)$ are equal when $i+j = \hat{x}_0 f_0$. However, for $2N < \hat{x}_0 - f_0$ the rates are all unique, and the probability density function is as follows:

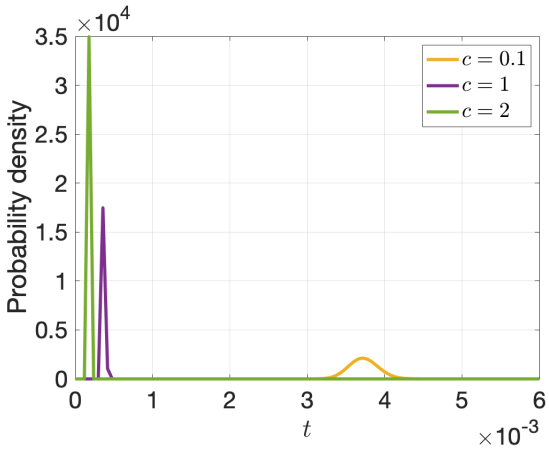


Figure 4.23: Analytically-derived examples of distributions for the characteristic monomer depletion time $t_{1/10}$, for the model with branching but not filament extension, for different values of c, the rate of branching. As c increases, $t_{1/10}$ decreases.

$$
\begin{aligned}
f_{t_{N/\hat{x}_0}}(t) &= \sum_{i=1}^{N} c(f_0+i-1)(\hat{x}_0-i+1) \\
&\times e^{-c(f_0+i-1)(\hat{x}_0-i+1)t}\left(\prod_{j=1,j\neq i}^{N}\frac{c(f_0+i-1)(\hat{x}_0-j+1)}{c(f_0+i-1)(\hat{x}_0-j+1)-c(f_0+i-1)(\hat{x}_0-i+1)}\right) \\
&= c\frac{(f_0+N-1)!\,\hat{x}_0!}{(f_0-1)!\,(\hat{x}_0-N)!}\sum_{i=0}^{N-1}\frac{(\hat{x}_0-f_0-i-N)!}{i!\,(N-i-1)!\,(x_0-f_0-i)!}(-1)^i(\hat{x}_0-f_0-2i)e^{-ct(f_0+i)(\hat{x}_0-i)}
\end{aligned}
\tag{4.24}
$$

When the rates of a hypoexponential function are not unique, the pdf does not have a simple closed form, but can be calculated using the phase-type representation outlined previously. Example plots of the probability density function are given in Figure 4.23; notably, as c increases, the mean and variance of $t_{N/\hat{x}_0}$ decrease. In general, the means and variances of the monomer depletion characteristic times are considerably less than their equivalents in the model with filament extension only.

4.4 Discussion

4.4.1 Summary

In this chapter, I have established analytical solutions for the deterministic simplified model of inflammasome formation and, in particular, the distribution of cluster sizes at each time point. I have also established the distribution of the following characteristic times in the stochastic version of the simplified model: the distribution of times of first joining events,

the time of emergence of a single cluster, and the time taken for ASC monomer abundance to decrease by a fraction $1/q$. These solutions suggest a number of conclusions; firstly, from the deterministic model we can see that the abundance of free monomers decreases approximately exponentially over time towards zero, with the rate of decrease increasing with the rate of filament branching and extension, as well as the initial number of NLR oligomers. The absence of a 'lag' at the beginning of this process is notable, and will be examined in comparison to experimental data in Chapter 6. The rate of increase of the mean number of ASC filaments and ASC monomers in each cluster levels off after an initial period of fast expansion in many cases. The rate of increase in cluster size is particularly marked when branching reactions are included, since these lead to a 'positive feedback loop' in which the creation of filaments also increases the overall rate of cluster growth. A similar effect can be observed for the variances of these variables. However, there are limits to the conclusions we can draw from the model at later times. Asymptotic analysis demonstrates that for all parametrisations of the model, the mean and variance of cluster sizes increase indefinitely over time, which is at odds with the depletion of ASC monomers and the emergence of an inflammasome with static size which is observed experimentally. This is caused by the inability of the model to accurately capture the system when there are small numbers of reactants.

I have also derived the distributions of the first joining time, the single cluster emergence time and the monomer depletion characteristic times from the stochastic formulation of the simplified model. A particularly notable feature of these results is that both joining rate of clusters and the initial number of NLR oligomers affect the time of first joining (increasing either parameter decreases the mean and variance of the first joining time); however, distribution of the time of single cluster emergence is mainly determined by the rate of cluster joining, and not the initial number of NLR oligomers. This suggests that the variation between cells in terms of inflammasome formation time may be due to inherent variation in cluster joining rate between cells. This could be caused, for example, by cells being different sizes, or containing varying levels of cellular crowding. However, an intrinsic variation in cluster joining rate would not explain the fact that in some inflammasome systems (for example, NLRP3 inflammasomes) not all cells form inflammasomes, whereas this is not the case, for example, with NLRC4 inflammasome formation. However, I will re-examine this conclusion in Chapter 5 in the light of results from the full model.

4.4.2 Further work

A natural next step is compare the results from the simplified to results from the full model; in particular, the different structure of the joining kernel in the full model is likely to lead to different results for characteristic times related to cluster joining (i.e. first joining time and single cluster emergence time). This will be considered in Chapter 5.

The distributions for characteristic times can also be compared with those derived from experimental data; this will be considered in Chapter 6, and fitting will be carried out using maximum likelihood estimation. One drawback of the analysis carried out here is that the general expression for the time of single cluster emergence is very unwieldy, and relies on taking the matrix exponential of a generally very large transition matrix. While this expression can be used to calculate a probability density function given all parameter values, this process takes too long to be useful when using a computational algorithm to minimise the resultant likelihood functions. I have been as yet unable to find a simple closed form for the probability density function for this very structured phase-type distribution, but this may be possible.

It may also be possible to extend the analysis carried out in this chapter, in order to establish the distribution of the times of dominant cluster emergence. This would be more difficult to derive than the single cluster emergence time, since this would have to take cluster size into account; however, for example, mean cluster size at a given time could be derived by calculating the distribution of the number of free ASC monomers by considering the Markov chain presented in Figure 4.17, and use conservation of monomers to derive the number of monomers in clusters; this could be divided by the distribution of the number of clusters which can be calculated from the Markov chain given in Figure 4.13. This would require considerable further analytical work, and it would be unlikely that a simple closed form could be found.

Another possible application is to extend the techniques described in this chapter to the full model; although this would be considerably more complex than the analysis carried out so far, it would be possible, for example, to define a phase-type distribution for the time of single cluster emergence in the full model. The generator matrix would be considerably larger, and would be unlikely to be upper triangular; due to the inclusion of movements between compartments, the monomer abundance is not monotonically decreasing, and a single ordering of states could not be contrived such that all movement between states occurs

in a single direction. While the generator matrix would still be sparse, calculations involving probability density functions and mean and variance would be significantly slower than those required for a triangular matrix.

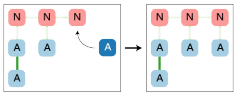
Chapter 5

Analysis and simulations of the full model of inflammasome formation

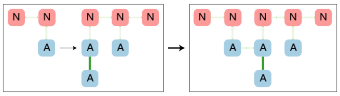
5.1 Introduction

Having analysed the simplified inflammasome model, we will now return to the full model of inflammasome formation initially proposed in Chapter 3 (Figure 5.1). The two compartments corresponding to the nucleus and cytosol are reintroduced, as is the distinction between unbound NLRs in oligomers and growing ASC filaments. The rate of cluster joining is also adjusted so that it is once more proportional to cluster size.

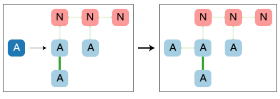
I will first examine the behaviour of the deterministic ODE formulation of the full inflammasome model; in particular, I will demonstrate the existence of an initial 'lag' phase in cluster



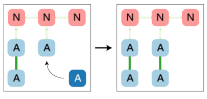
Reaction C1 (Recruitment of ASC monomer by NLR in oligomer)



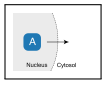
Reaction C2 (Cluster joining)



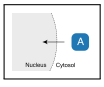
Reaction C3 (Branching of ASC filaments)



Reaction P (ASC filament extension)



Reaction M1 (Movement of ASC monomer from nucleus to cytosol)



Reaction M2 (Movement of ASC monomer from nucleus to cytosol)

Figure 5.1: The full inflammasome model revisited

growth and monomer depletion when ASC recruitment rate by NLRs is slow; this is a clear contrast between the full and simplified inflammasome models. I will also define a set of parametrisations for the model for which the lag phase does not exist. However, as I did for the simplified model, I will also demonstrate that the full ODE model displays non-physical behaviour at later timepoints as the concentration of reactants becomes smaller.

I will then present results from simulations of the stochastic formulation of the full model. In particular, I will demonstrate that the process of inflammasome formation up to single cluster emergence can be divided into separate phases: an initial phase before cluster joining commences, an intermediate phase in which clusters begin to join, and a later phase in which a single dominant cluster emerges. I will also demonstrate that the variation between simulation runs (and by analogy, between individual cells) arises from stochasticity in the second phase.

Finally, I will present results from a cohort of simulations run with parameter sets chosen using Latin hypercube sampling. In order to determine the dependence of the characteristic times for each simulation run on the parameter set used, I will fit a generalised linear model to the results for each characteristic time. I will use this model to demonstrate that while early characteristic times are dependent on a large set of parameters, later characteristic times (namely, the times of single and dominant cluster emergence) are dependent on very few parameters. In particular, the time of dominant cluster emergence is dependent on the branching rate and the initial number of NLR oligomers, whereas the time of single cluster emergence is dependent on the branching rate and the rate of movement of ASC monomers out of the nucleus.

I have also generated simulations for variations on the full inflammasome model, in which some of the initial assumptions have been altered. In particular, I have considered a variation on the full model in which there is no filament branching, a model in which there is only one NLR oligomer initially present, and a model in which the initial number of NLRs in each oligomers is random rather than constant. I will demonstrate the differences arising between these models and the full models with basic assumptions; all four datasets will be compared to experimental data in Chapter 6.

5.2 Analysis of the ODE model

We will begin by considering the deterministic formulation of the full inflammasome model using the reduced ODE form:

$$
\begin{aligned}
\frac{\mathrm{d}\hat{x}_c}{\mathrm{d}t} &= -c_1\hat{x}_c N - c_3\hat{x}_c F - p\hat{x}_c F + m_1\hat{x}_n - m_2\hat{x}_c \\
\frac{\mathrm{d}\hat{x}_n}{\mathrm{d}t} &= -m_1\hat{x}_n + m_2\hat{x}_c \\
\frac{\mathrm{d}X}{\mathrm{d}t} &= -\frac{1}{2}c_2F^2 \\
\frac{\mathrm{d}F}{\mathrm{d}t} &= c_1\hat{x}_c N + c_3\hat{x}_c F \\
\frac{\mathrm{d}N}{\mathrm{d}t} &= -c_1\hat{x}_c N \\
\frac{\mathrm{d}A}{\mathrm{d}t} &= c_1\hat{x}_c N + c_3\hat{x}_c F + p\hat{x}_c F \\
\frac{\mathrm{d}s_{F,F}}{\mathrm{d}t} &= c_1\hat{x}_c(2s_{F,N} + N) + c_2 s_{F,F}^2 + c_3\hat{x}_c(2s_{F,F} + F) \\
\frac{\mathrm{d}s_{N,N}}{\mathrm{d}t} &= c_1\hat{x}_c(N - 2s_{N,N}) + c_2 s_{F,N}^2 \\
\frac{\mathrm{d}s_{A,A}}{\mathrm{d}t} &= c_1\hat{x}_c(2s_{N,A} + N) + c_2 s_{F,A}^2 + (c_3 + p)\hat{x}_c(2s_{F,A} + F) \\
\frac{\mathrm{d}s_{F,N}}{\mathrm{d}t} &= c_1\hat{x}_c(s_{F,F} - s_{F,N} - N) + c_2 s_{F,F}s_{F,N} + c_3\hat{x}_c s_{F,N} \\
\frac{\mathrm{d}s_{F,A}}{\mathrm{d}t} &= c_1\hat{x}_c(s_{N,A} + s_{F,N} + N) + c_2 s_{F,F}s_{F,A} + c_3\hat{x}_c(s_{F,F} + s_{F,A} + F) + p\hat{x}_c s_{F,F} \\
\frac{\mathrm{d}s_{N,A}}{\mathrm{d}t} &= c_1\hat{x}_c(s_{N,N} - s_{N,A} - N) + c_2 s_{F,N}s_{F,A} + (c_3 + p)\hat{x}_c s_{F,N}
\end{aligned}
\tag{3.2 revisited}
$$

Here, $\hat{x}_c$ and $\hat{x}_n$ are the concentrations of ASC monomers in the cytosolic and nuclear compartments respectively. X is the total concentration of clusters, F is the total concentration of filaments in the system, N is the total concentration of unbound NLRs, and A is the total concentration of ASC monomers in clusters. The variables $S_{I,J}$ are the second moments of the system, which can be used to calculate the centralised moments:

$$\mu_F = \frac{F}{X} \tag{5.1a}$$

$$\sigma_F^2 = \frac{S_{F,F}}{X} - \mu_F^2 \tag{5.1b}$$

$$\mu_A = \frac{A}{X} \tag{5.1c}$$

$$\sigma_A^2 = \frac{S_{A,A}}{X} - \mu_A^2 \tag{5.1d}$$

Here, as for the simplified model in Chapter 4, μ_F and μ_A are the mean number of

$$\begin{aligned} \hat{x}_c(0) &= \hat{x}_0^{(c)} & s_{F,F}(0) &= 0 \\ \hat{x}_n(0) &= \hat{x}_0^{(n)}) & s_{N,N}(0) &= 100X_0^2 \\ X(0) &= X_0 & s_{A,A} &= 0 \\ F(0) &= 0 & s_{F,N} &= 0 \\ N(0) &= 10X_0 & s_{F,A} &= 0 \\ A(0) &= 0 & s_{N,A} &= 0 \end{aligned}$$

filaments and ASC monomers per cluster respectively, and σ_F^2 and σ_A^2 are the variances of the number of filaments and ASC monomers per cluster respectively.

The intial conditions for the system of equations 3.2 are as follows:

Although (unlike the simplified ODE model) System 3.2 is not analytically solvable, numerical solutions are useful for comparison; I have implemented the system using the `ode45` solver. Values of μ_F, μ_n,μ_A, σ_F^2, σ_N^2 and σ_A^2 can then be calculated.

Throughout my analysis of the full deterministic model, I set $m_2 = 1.3 \times m_1$, and thus $\hat{x}_n(0) = 1.3 \times \hat{x}_c(0)$, assuming the populations of ASC monomers begin in equilibrium. This value is estimated from experimental data in Section 2.4.1. As in the analysis of the simplified model, I set $c_2 = 1$ and $\hat{x}_c(0) = 1$ throughout, and asssumed that all clusters begin with 10 NLRs.

One of the notable properties of the full ODE model is that the concentration of clusters X eventually becomes negative; this is clearly non-physical, and certainly undesirable. This situation arises due to the size-dependent joining process. In particular, the evolution of X is governed by the equation $\frac{\mathrm{d}X}{\mathrm{d}t} = -\frac{1}{2}c_2F^2$. However, the system clearly tends to a fixed point where $\hat{x}_c = \hat{x}_n = N = 0$, and F tends towards a positive constant value. Therefore, at some point X must become negative. This issue does not arise in the simplified model, since the rate of cluster joining is dependent only on the number of clusters; in particular, $\frac{\mathrm{d}X}{\mathrm{d}t} = -\frac{1}{2}c_2X^2$, so we certainly have $X \to 0$ as $t \to \infty$.

This non-physical behaviour arises from the approximation made in the joining term while setting up the ODE model, in which we do not prevent clusters from joining with themselves (this is elaborated in Section 3.2.4); essentially X eventually becomes negative, because when X is small the majority of filaments in fact belong to the same cluster, so there are far fewer possible joining events than the ODE model suggests. This is an unavoidable feature within the continuous framework in which $\hat{x}_c$, $\hat{x}_n$ and other variables represent concentration

rather than absolute number of reactants. In order to gain a truly accurate picture of the behaviour of the system when there are small numbers of reactants, a discrete stochastic system is more appropriate.

On the other hand, the issue of undesirable behaviour of the ODE model is not present when reactant numbers are large; therefore it is worth considering the ODE model to understand the behaviour at early times, until the point where we have $X < 0$ and the results become non-physical.

5.2.1 Dynamics of ASC monomer concentration

In the full model, at late times $\hat{x}_c$ and $\hat{x}_n$ have similar dynamics (exponential decreases towards zero) to the simplified model for m_1 and m_2 larger than the cluster growth and joining parameters (Figures 5.2 and 5.3). However, depletion is slightly slower for $\hat{x}_n$, which is also to be expected, since ASC monomers are only depleted by cluster growth in the cytosolic compartment. However, in the case of small c_1, when binding of monomers to NLRs are limited, there is a notable lag before the exponential decrease begins, leading to an overall sigmoid curve; this is due to an initial bottleneck in cluster growth as filament seeding is slow (Figures 5.2 and 5.3). The smaller the value of c_1, the longer this lag phase. This is in contrast to the simplified model, in which this bottleneck is absent and no initial lag occurs.

In general, the rates of monomer depletion increase with increasing c_1, c_3 and p, and slightly with increasing m_1. The dynamics of $\hat{x}_c$ and $\hat{x}_n$ are not affected by X_0, since the depletion of monomers is dependent only on the total filaments present, regardless of the number of clusters.

5.2.2 Dynamics of total cluster concentration

The total cluster concentration similarly displays a decrease following initial lag phase, not present in the simplified model (Figure 5.4). This lag phase increases with decreasing c_1, once again due to the bottleneck in filament seeding; since in this model, clusters merge via filament-filament interactions, this cannot occur until sufficiently many filaments have been seeded. The lag time also increases with X_0, since it will therefore take longer for the same proportion of NLRs present to seed filaments. The rate of decrease of X decreases with p and increases with c_3 if all other variables are fixed; since cluster merging is filament-driven, filament branching contributes to faster merging, whereas linear filament growth does not.

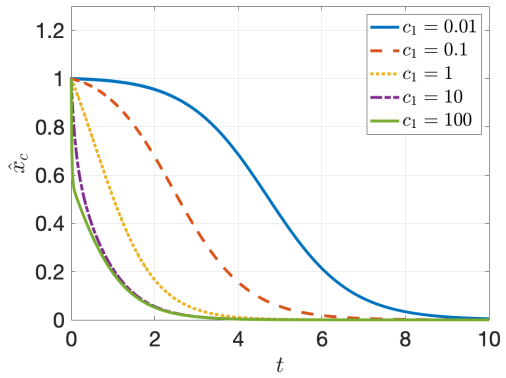
5.2.3 Distributions of cluster size

The dynamics of μ_F, μ_A, σ_F^2 and σ_A^2 differ considerably from their dynamics in the simplified model, bearing in mind that μ_F and σ_F^2 are not direct parallels in the simplified and full model, since in the simplified model, F also includes unbound NLRs (Figures 5.5 - 5.8). In the full model, all of these variables tend towards a sharp increase (much steeper than in the simplified model), following an initial lag phase with a slow increase not present in the simplified model. As in the simplified model, the average cluster size naturally increases as merging and growth occur, and the variance of cluster sizes increases as growth is dominated by a few larger clusters. However, the initial rate of growth is limited by the filament seeding bottleneck. Unlike the simplified model, in which an initial period of rapid growth gives way to linear growth (for mean cluster size variables) or quadratic growth (for variance variables), in the full model there is an initial slower period of growth which gives way to a sharp increase. The length of the initial lag phases increase with increasing X_0 and decrease with increasing c_1. This suggests the presence of separate phases in the cluster formation process: an initial phase of filament number increase in which linear filament growth dominates, followed by a phase of rapid mean filament number growth as clusters merge. However, in order to investigate this phenomenon further, we will turn to a stochastic formulation of the full model, which remains accurate at late times.

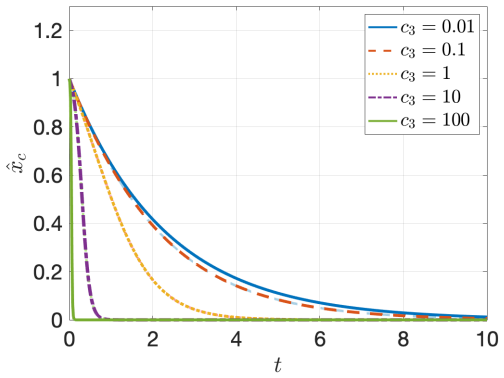
5.2.4 Identifying the length of lag phase for different parametrisations

In order to establish a relationship between the parametrisation of the model and the presence or absence of an initial sigmoid lag, I solved System 3.2 numerically with parameters c_1, c_3 and p ranging between 10^{-1} to 10, and X_0 ranging from 10^{-10} to 1. I took the calculated values for $\hat{x}_c$ and used an adaptation of a standard technique for determining lag phase in bacterial growth to gain an approximate measurement for the lag phase (illustrated in Figure 5.9; see, for example, the method outlined in [Adkar et al., 2017]).

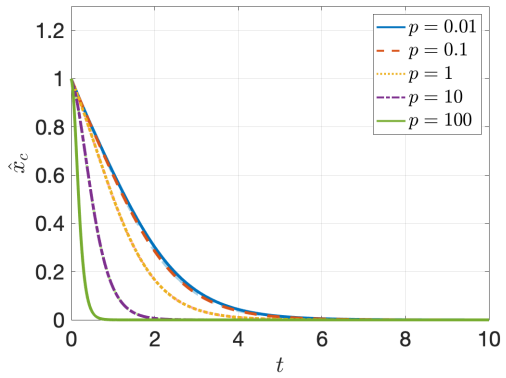
Overall, X_0 has little effect on the length of the lag phase, as observed previously; however, there is a clear set of values of the parameters c_1, c_3 and p for which the calculated lag phase has length zero - i.e., there is no lag before the rapid phase of ASC monomer depletion (Figure 5.10). For each pair of values c_3 and p, there is a maximum value of c_1 for which a lag phase occurs; above this value there is no lag phase (Figure 5.11). Using the



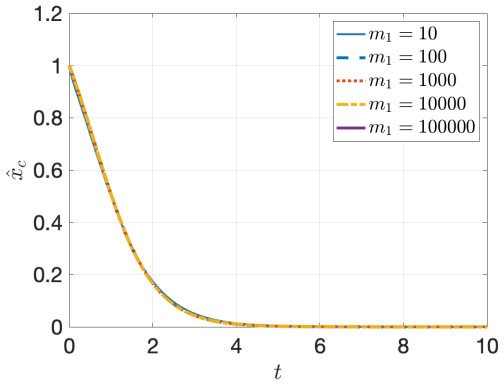
(a) $\hat{x}_c$ varying c_1



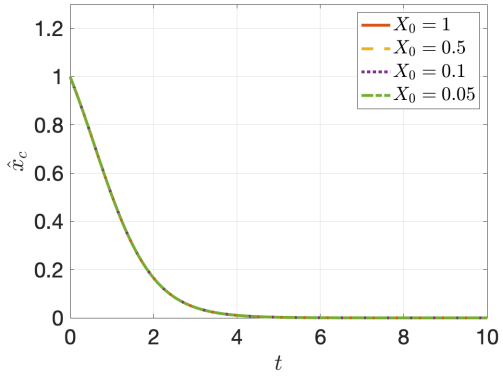
(b) $\hat{x}_c$ varying c_3



(c) $\hat{x}_c$ varying p

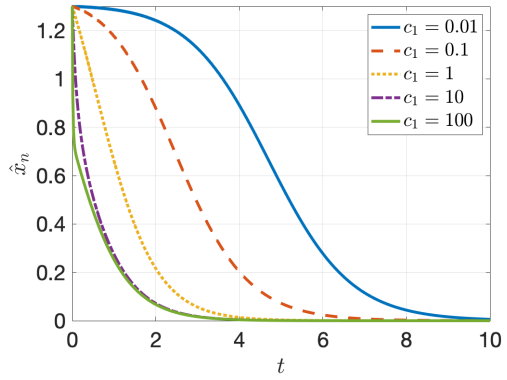


(d) $\hat{x}_c$ varying m_1

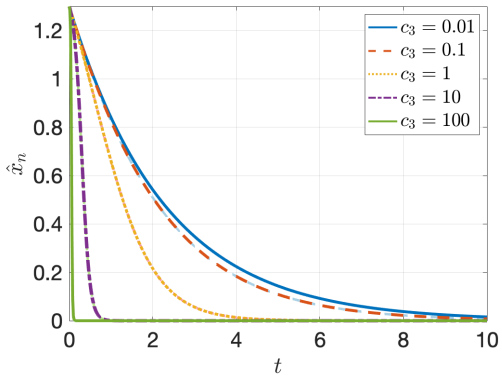


(e) $\hat{x}_c$ varying X_0

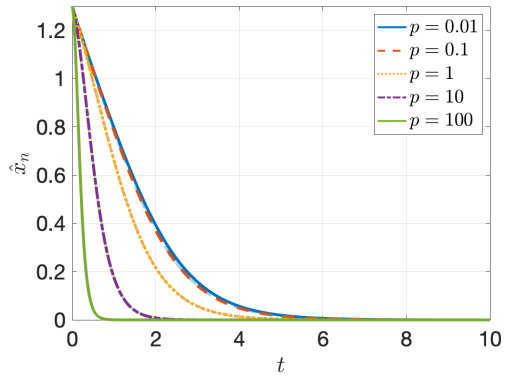
Figure 5.2: Cytosolic ASC concentration $\hat{x}_c$ over time while varying the rate of ASC recruitment by NLRs c_1, filament branching rate c_3, filament extension rate p, rate of movement of ASC out of the nucleus m_1 and initial NLR oligomer concentrationX_0. Numerical solutions derived from the ODE system are shown. Cluster joining rate $c_2 = 1$ and initial monomer concentration $\hat{x}_c(0) = 1$ are fixed; $c_1 = c_3 = 1, m_1 = 1000$ and $X_0 = 0.1$ are the default values when not being varied for comparison. Monomers are depleted in a sigmoid fashion. In the case of small c_1, there is a notable lag phase; the smaller the value of c_1, the longer this lag phase. Rates of monomer depletion increase with increasing c_1, c_3 and p, and slightly with increasing m_1. Monomer dynamics are not affected by the value of X_0.



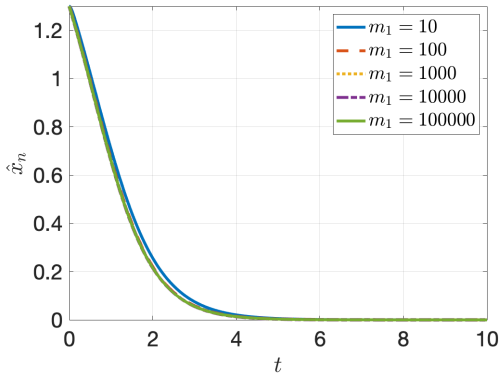
(a) $\hat{x}_n$ varying c_1



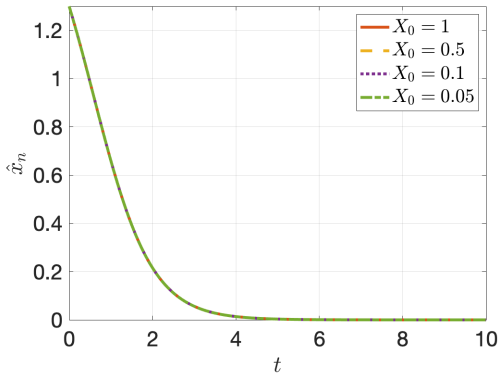
(b) $\hat{x}_n$ varying c_3



(c) $\hat{x}_n$ varying p

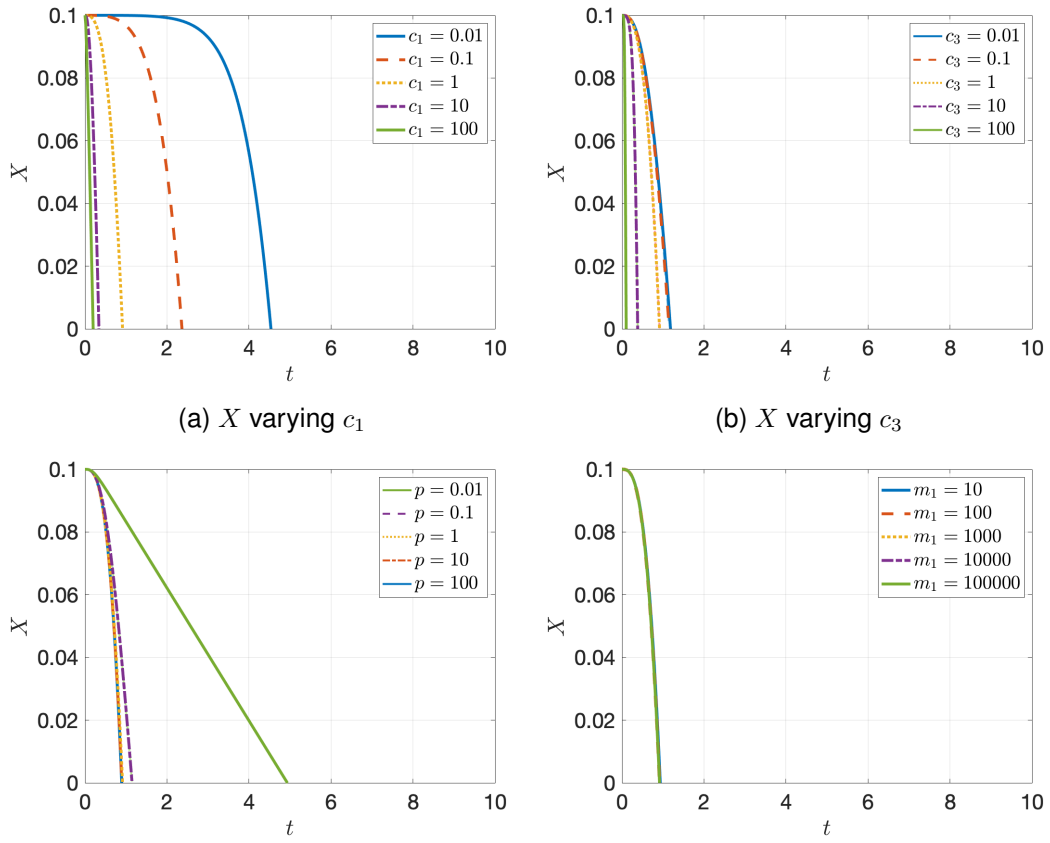


(d) $\hat{x}_n$ varying m_1



(e) $\hat{x}_n$ varying X_0

Figure 5.3: Nuclear ASC concentration over time while varying the rate of ASC recruitment by NLRs c_1, filament branching rate c_3, filament extension rate p, rate of movement of ASC out of the nucleus m_1 and initial NLR oligomer concentration X_0. Numerical solutions derived from the ODE system are shown. Cluster joining rate $c_2 = 1$ and initial monomer concentration $\hat{x}_c(0) = 1$ are fixed; $c_1 = c_3 = 1, m_1 = 1000$ and $X_0 = 0.1$ are the default values when not being varied for comparison. Overall, dynamics are very similar to those for $\hat{x}_c$, albeit slightly slower.



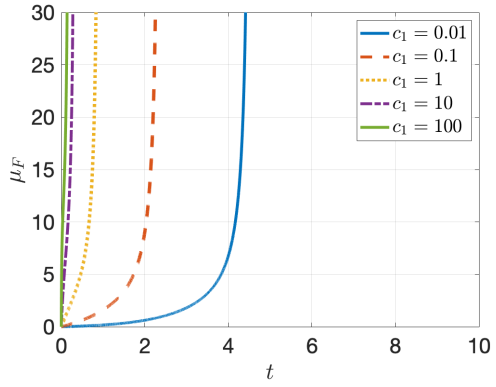
(a) X varying c_1

(b) X varying c_3

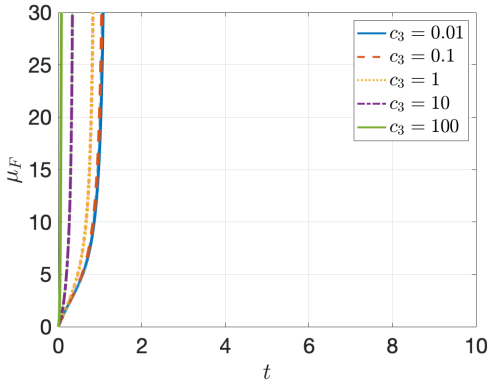
(c) X varying p

(d) X varying m_1

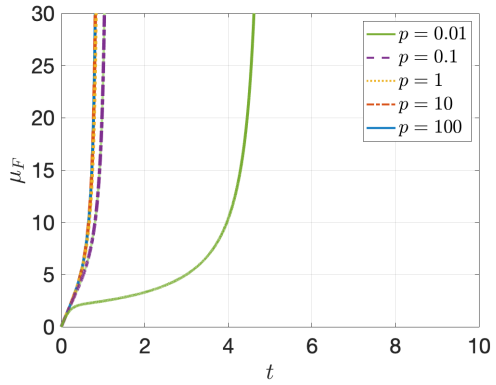
Figure 5.4: Cluster concentration X over time while varying the rate of ASC recruitment by NLRs c_1, filament branching rate c_3, filament extension rate p, rate of movement of ASC out of the nucleus m_1 and initial NLR oligomer concentration X_0. Numerical solutions derived from the ODE system are shown. Cluster joining rate $c_2 = 1$ and initial monomer concentration $\hat{x}_c(0) = 1$ are fixed; $c_1 = c_3 = 1, m_1 = 1000$ and $X_0 = 0.1$ are the default values when not being varied for comparison. X decreases following an initial lag phase; the length of the lag phase increases with decreasing c_1 and with increasing X_0. The rate of decrease of X decreases with p and increases with c_3. Varying m_1 has little effect on the dynamics of X. In all cases, X ultimately crosses the x-axis, and takes non-physical negative values; after this point, the model is no longer biologically plausible.



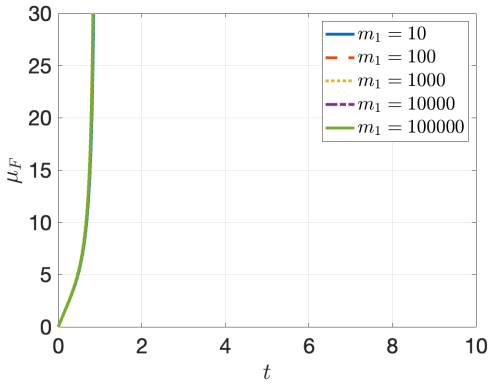
(a) μ_F varying c_1



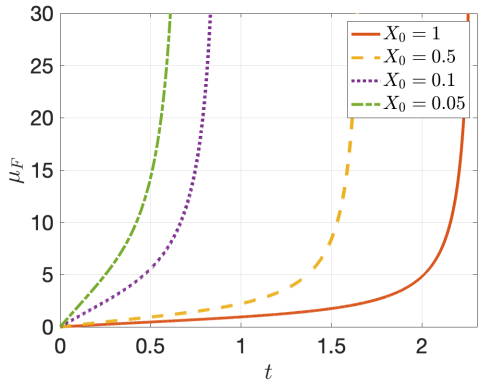
(b) μ_F varying c_3



(c) μ_F varying p

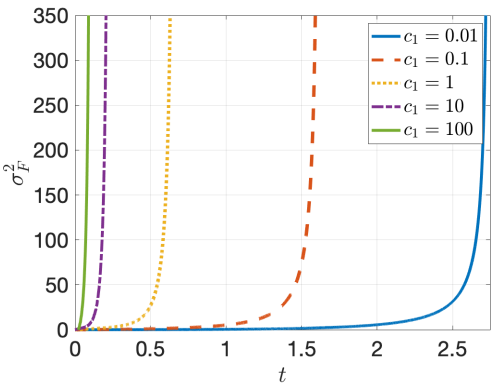


(d) μ_F varying m_1

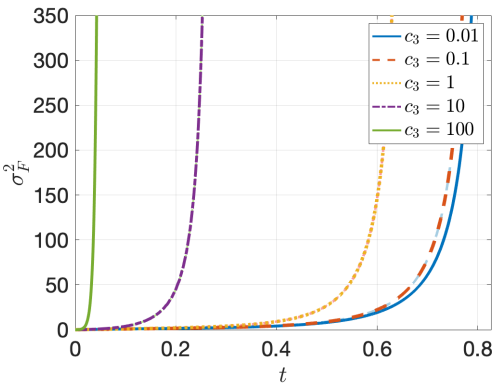


(e) μ_F varying X_0

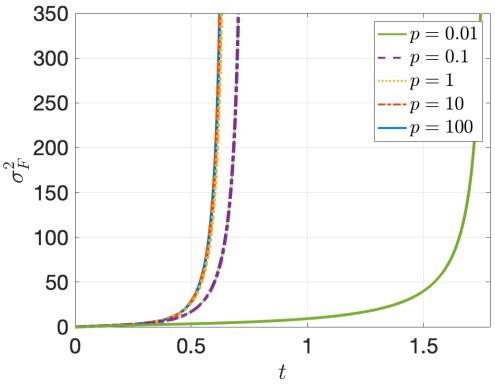
Figure 5.5: Mean number of filaments per cluster over time while varying the rate of ASC recruitment by NLRs c_1, filament branching rate c_3, filament extension rate p, rate of movement of ASC out of the nucleus m_1 and initial NLR oligomer concentration X_0. Numerical solutions derived from the ODE system are shown. Cluster joining rate $c_2 = 1$ and initial monomer concentration $\hat{x}_c(0) = 1$ are fixed; $c_1 = c_3 = 1, m_1 = 1000$ and $X_0 = 0.1$ are the default values when not being varied for comparison. μ_F increases rapidly following a lag phase; the length of lag increases with increasing X_0 and decreases with increasing c_1.The rate of increase at late times increases with increasing c_3.



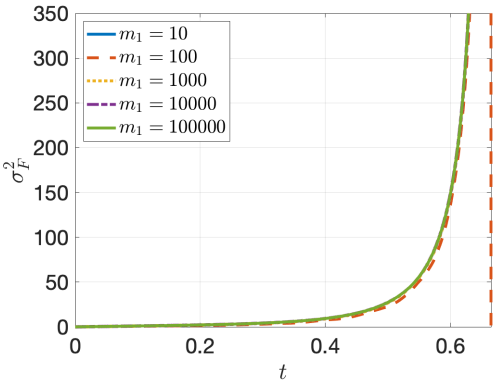
(a) σ_F^2 varying c_1



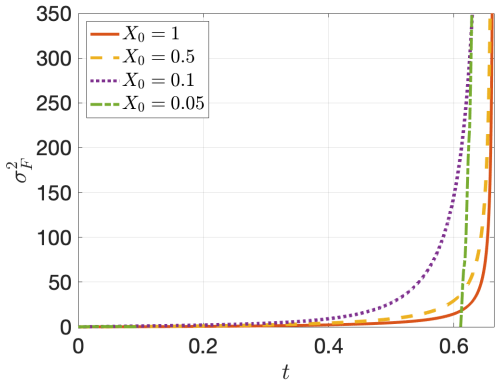
(b) σ_F^2 varying c_3



(c) σ_F^2 varying p

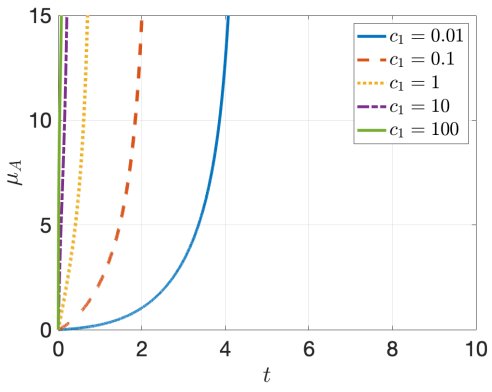


(d) σ_F^2 varying m_1

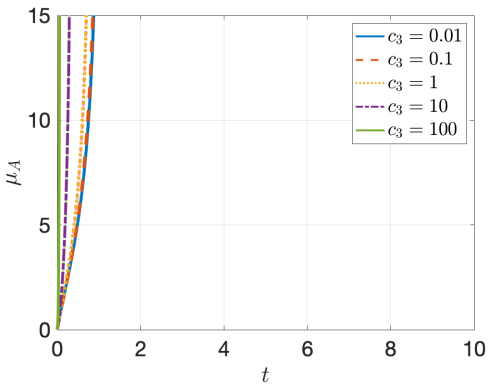


(e) σ_F^2 varying X_0

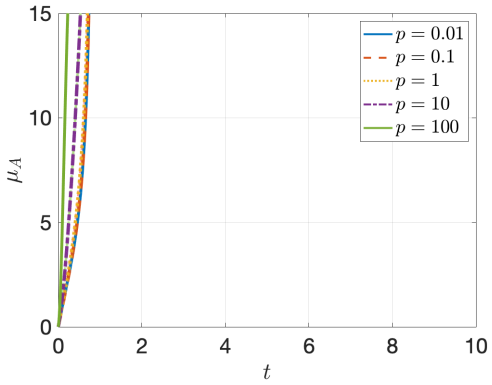
Figure 5.6: Variance of number of filaments per cluster over time while varying the rate of ASC recruitment by NLRs c_1, filament branching rate c_3, filament extension rate p, rate of movement of ASC out of the nucleus m_1 and initial NLR oligomer concentration X_0. Numerical solutions derived from the ODE system are shown. Cluster joining rate $c_2 = 1$ and initial monomer concentration $\hat{x}_c(0) = 1$ are fixed; $c_1 = c_3 = 1, m_1 = 1000$ and $X_0 = 0.1$ are the default values when not being varied for comparison. σ_F^2 increases rapidly following a lag phase; the length of lag increases with increasing X_0 and decreases with increasing c_1.



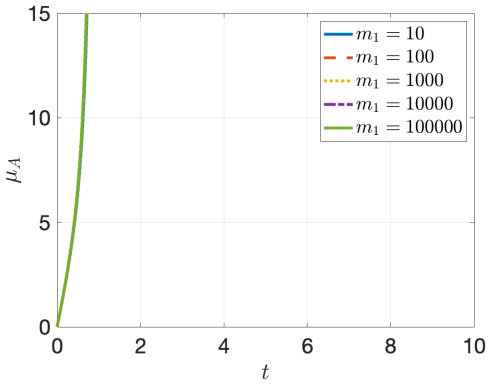
(a) μ_A varying c_1



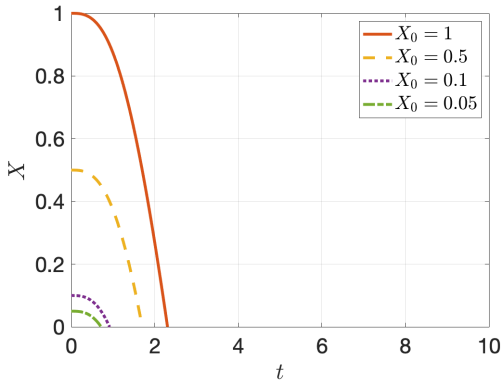
(b) μ_A varying c_3



(c) μ_A varying p

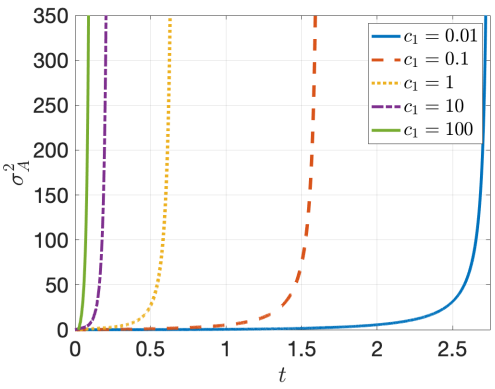


(d) μ_A varying m_1

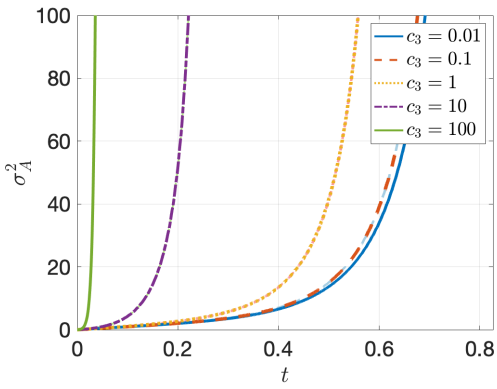


(e) μ_A varying X_0

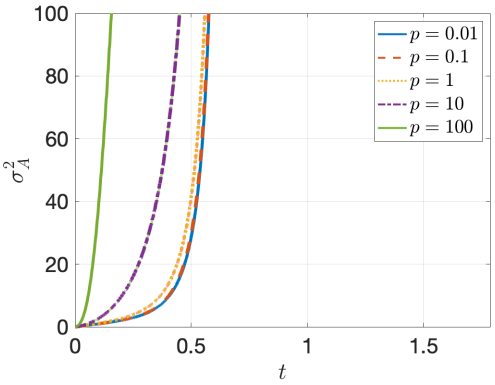
Figure 5.7: Mean ASC monomers per cluster μ_A over time while varying the rate of ASC recruitment by NLRs c_1, filament branching rate c_3, filament extension rate p, rate of movement of ASC out of the nucleus m_1 and initial NLR oligomer concentration X_0. Numerical solutions derived from the ODE system are shown. Cluster joining rate $c_2 = 1$ and initial monomer concentration $\hat{x}_c(0) = 1$ are fixed; $c_1 = c_3 = 1, m_1 = 1000$ and $X_0 = 0.1$ are the default values when not being varied for comparison. μ_A increases rapidly following a lag phase; the length of lag increases with increasing X_0 and decreases with increasing c_1.



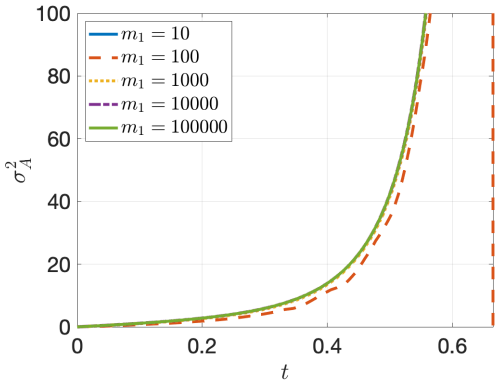
(a) σ_A^2 varying c_1



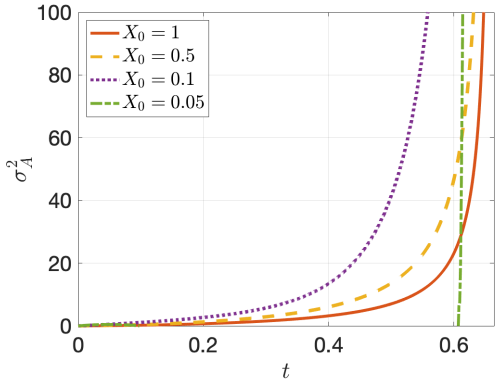
(b) σ_A^2 varying c_3



(c) σ_A^2 varying p



(d) σ_A^2 varying m_1



(e) σ_A^2 varying X_0

Figure 5.8: Variance of ASC monomers per cluster σ_A^2 over time while varying the rate of ASC recruitment by NLRs c_1, filament branching rate c_3, filament extension rate p, rate of movement of ASC out of the nucleus m_1 and initial NLR oligomer concentration X_0. Numerical solutions derived from the ODE system are shown. Cluster joining rate $c_2 = 1$ and initial monomer concentration $\hat{x}_c(0) = 1$ are fixed; $c_1 = c_3 = 1, m_1 = 1000$ and $X_0 = 0.1$ are the default values when not being varied for comparison. σ_A^2 increases rapidly following a lag phase; the length of lag increases with increasing X_0 and decreases with increasing c_1.

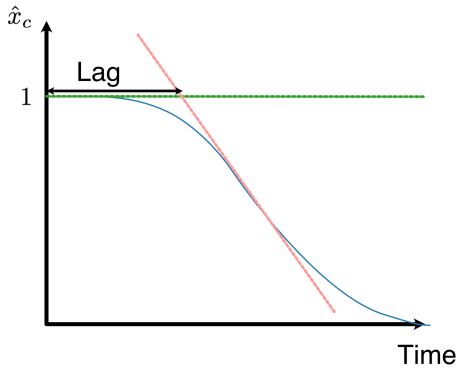


Figure 5.9: Schematic diagram showing calculation of the lag phase at the beginning of a cytosolic ASC monomer trace (shown in blue). The point of inflection is found, and a tangent line drawn at this point (shown in pink). The end of the lag phase is taken as the intercept of the tangent with the line $\hat{x}_c = 1$.

Reference	**Reaction**	**Propensity function**
C1	$\hat{x}_c + x_{f,n,a} \xrightarrow{c_1 n} x_{f+1,n-1,a+1}$	$c_1 n \hat{x}_c x_{f,n,a}$
C2	$x_{f,n,a} + x_{i,j,k} \xrightarrow{c_2 fi} x_{f+i,n+j,a+k}$	$c_2 fi x_{f,n,a} x_{i,j,k}$ for $(f,n,a) \neq (i,j,k)$, $\frac{1}{2} c_2 f^2 x_{f,n,a}(x_{f,n,a} - 1)$ otherwise
C3	$\hat{x}_c + x_{f,n,a} \xrightarrow{c_3 f} x_{f+1,n,a+1}$	$c_3 f \hat{x}_c x_{f,n,a}$
P	$\hat{x}_c + x_{f,n,a} \xrightarrow{pf} x_{f,n,a+1}$	$pf \hat{x}_c x_{f,n,a}$
M1	$\hat{x}_n \xrightarrow{m_1} \hat{x}_c$	$m_1 \hat{x}_n$
M2	$\hat{x}_c \xrightarrow{m_2} \hat{x}_n$	$m_2 \hat{x}_c$

Table 5.1: Rates of reactions in the stochastic formulation of the full inflammasome model

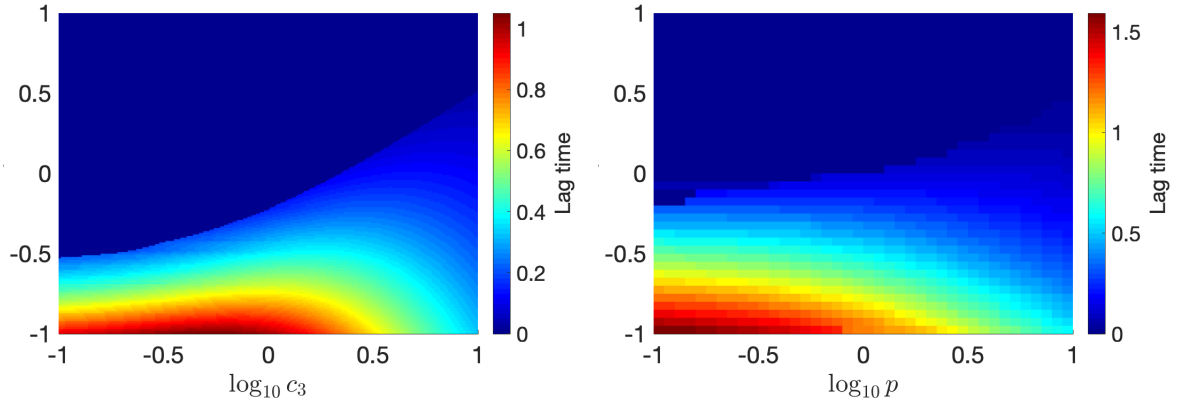
Matlab function `fit` to fit a quadratic in c_3 and p, an approximation for the surface defined by these maximum values of c_1 is as follows:

$$c_1 > 0.546 + 0.385 c_3 + 0.209 p + 3.60 \times 10^{-3} c_3^2 - 3.46 \times 10^{-3} c_3 p + 9.48 \times 10^{-4} p^2 \quad (5.2)$$

This approximately defines the regions in the parameter space for which a lag phase appears in $\hat{x}_c$; this will be used in fitting data in Chapter 6.

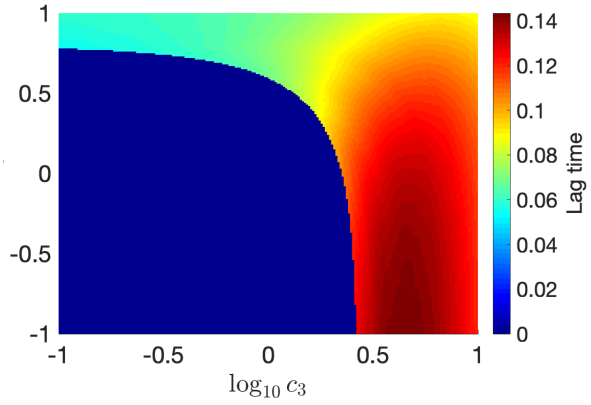
5.3 Simulations of the stochastic model

I will now turn my attention to the stochastic formulation of the full inflammasome model. The propensity functions, as first presented in Chapter 3, are given in Table 5.1.

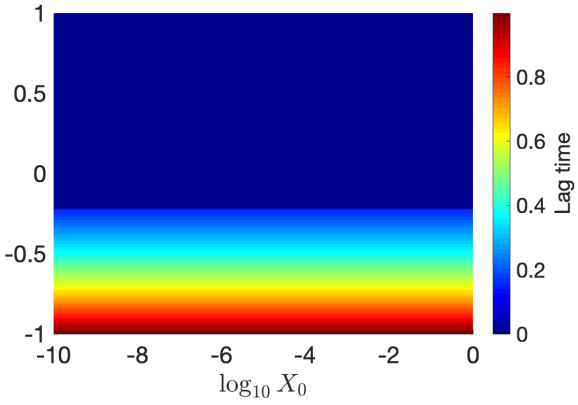


(a) Varying c_1 and c_3

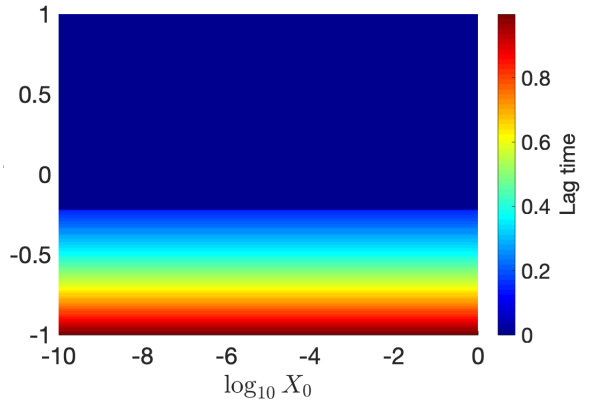
(b) Varying c_1 and p



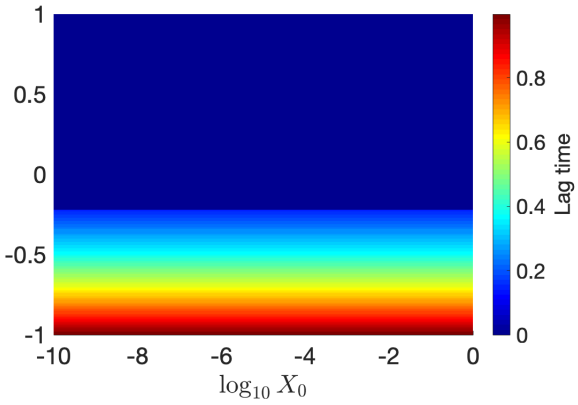
(c) Varying c_3 and p



(d) Varying c_1 and X_0



(e) Varying c_3 and X_0



(f) Varying p and X_0

Figure 5.10: Length of lag phase for the ASC monomer concentration in the cytosolic compartment varying the rate of ASC recruitment by NLRs, along with the branching rate c_3, the filament extension rate p and the initial number of oligomers X_0. Results have been obtained from numerical solutions of the ODE model. Unless values are shown, $c_3 = 1, p = 1, m_1 = 1000, m_2 = 1.3m_1, X_0 = 0.1$.

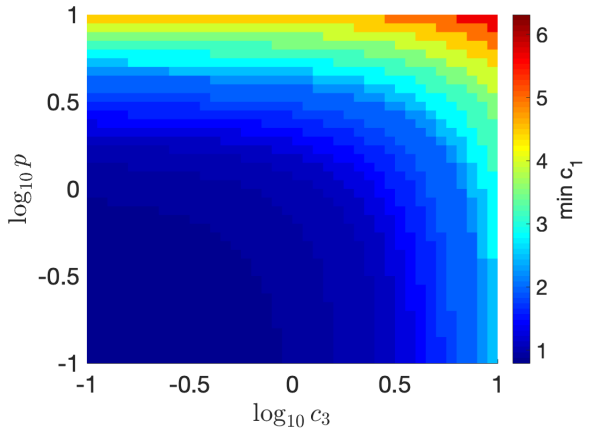


Figure 5.11: Minimum values of the rate of ASC recruitment by NLRs c_1 for which there is no lag phase, for a range of values of the branching rate c_3 and the filament extension rate p. For values of c_1 beneath this minimum value, there is a lag phase of nonzero length. Results have been obtained from numerical solutions of the ODE model. The minimum value of c_1 for which there is no lag phase is higher for larger c_3 and p.

5.3.1 Simulation methods

I have implemented the discrete stochastic model in MATLAB using a variation on the Gillespie stochastic simulation algorithm [Gillespie, 1976]. Since I have assumed the movement of ASC monomers between cellular compartments is much faster than cluster growth and joining reactions, it becomes very inefficient to use the Gillespie algorithm to simulate the cluster joining and growth reactions and monomer movement reactions simultaneously. Therefore, I use the Gillespie algorithm to simulate cluster joining and growth reactions C1, C2, C3 and P only. I assume that the number of ASC monomers in the nuclear and cytosolic compartments are approximately constant between reaction steps; this is justified by the fast timescale of these reactions. The abundance of ASC monomers in each compartment is updated at each timestep generated by the Gillespie algorithm, unless this timestep is longer than a maximum time δt, in which case, the complex joining or growth reaction is discarded, and the number of ASC monomers in each compartment is updated after a timestep of δt. The Gillespie algorithm is then resumed. The change in number of ASC monomers in each compartment is calculated by assuming the movement of monomers out of a given compartment is binomially distributed; specifically, the number of monomers moving from the nucleus to cytosol between times t_1 and t_2 is binomially distributed with $\hat{x}_n(t_1)$ trials and probability of success $\frac{m_1}{m_1+m_2}\left(1-e^{-(m_1+m_2)(t_2-t_1)}\right)$. The success probability arises from considering exponentially-distributed jump times of individual monomers. Similarly, the num-

ber of monomers moving from cytosol to nucleus between times t_1 and t_2 is binomially distributed with $\hat{x}_c(t_1)$ trials and success probability $\frac{m_2}{m_1+m_2}\left(1-e^{-(m_1+m_2)(t_2-t_1)}\right)$. A pair of binomially-distributed random numbers with these parameters is therefore generated for each time point arising in the simulation. Setting $\delta t = 10^{-5}$ was sufficient to avoid major simulation artefacts for parameters we investigated.

Note that in the discrete stochastic system, self-joining is not permitted; this adds a correction term to the rate of joining of two complexes of the same size (see Table 5.1.).

I ran simulations for the basic model outlined above, as well as model variants with only a single initial NLR oligomer (identical to the basic model with the initial number of NLR oligomers $X_0 = 1$. In order to test sensitivity to the assumption that each NLR oligomer contains exactly 10 monomers, for each value of X_0 I also ran simulations with stochastically generated initial conditions; in particular, instead of assuming that all oligomers contained exactly 10 NLRs, I distributed the $10X_0$ NLRs initially present randomly between the X_0 oligomers. I used the multinomial distribution to generate oligomer sizes, with $10X_0$ trials and all event probabilities $p_i = 1/X_0$. This mimics the assumption that NLR oligomers grow via a pool of NLR monomers being added at random to a fixed number (X_0) of growing oligomers until all monomers are depleted; since oligomer growth is linear, and independent of oligomer size. the assumption that the probability that a given NLR monomer joins a given oligomer is $1/X_0$ at all times is justified.

For each of these model variants, I varied $c_2, c_3, p, m_1, m_2/m_1$ and X_0 and fixed $c_1 = 1$ (since this parameter can be scaled out by rescaling time) and $\hat{x}_c(0) = 10^4$. In particular, I assumed that the number of ASC monomers in the nuclear and cytosolic compartments begin in equilibrium, so $\hat{x}_n(0) = \frac{m_2}{m_1} \times \hat{x}_c(0)$. In general, I have also taken the initial number of free ASC monomers in the cytosolic region, $\hat{x}_c(0)$, to be 10^4; once again this is somewhat arbitrary, as all abundances could be scaled out of the model.

Since the majority of the parameter values and ICs in the model are unknown, and in order to reduce the number of simulations needed to cover the parameter space, I carried out Latin hypercube sampling, as used by Proctor *et al.* for analysing NP models [Proctor et al., 2012]. I generated 100 sets of parameter values and ICs for each model variant, minimising correlation between parameters in the set; a summary of the ranges of values sampled for each parameter is given in Table 5.2. I have also given baseline values of each parameter, which I have used to create the example parameter runs depicted in the

Variable	Role in model	Maximum	Minimum	Baseline
c_2	Cluster joining rate	10^{-1}	10^{1}	1
c_3	Filament branching rate	10^{-1}	10^{1}	1
p	Filament extension rate	10^{-1}	10^{1}	1
m_1	Rate of movement of ASC monomers from nucleus to cytosol	10^{1}	10^{3}	10^{3}
m_2/m_1	Ratio of movement rates of ASC from nucleus to cytosol, and from cytosol to nucleus	0	2	1.5
X_0	Initial number of NLR oligomers	1	10^{3}	10^{2}

Table 5.2: Ranges of parameter values sampled and baseline values for running simulations of the full model of inflammasome formation

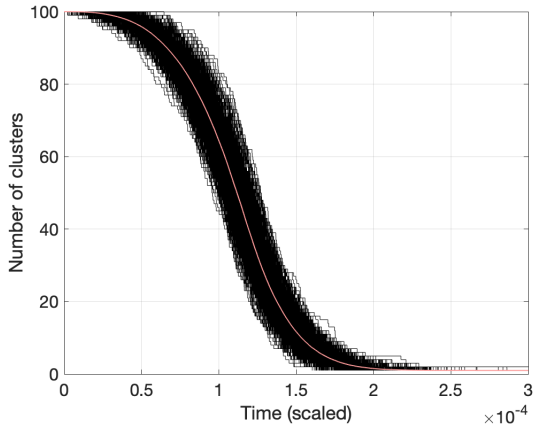
following section. For each parameter set, I ran 100 simulations, to ensure the robustness of the results. I also ran the simulations with baseline values 1000 times.

5.3.2 Dynamics of individual simulation runs with baseline parameter values

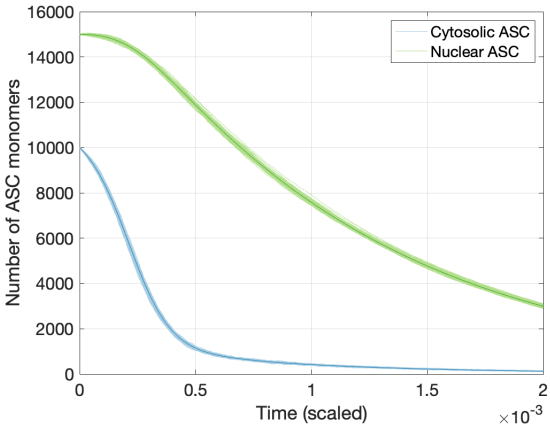
I will begin by considering results from the large set of simulations with baseline parameter values, in order to clarify the dynamics of individual simulation runs.

In general, as for the deterministic model results, the number of clusters takes a sigmoid shape over time, decreasing to one following an initial lag phase. Similarly, the number of monomers in nuclear and cytosolic compartments decreases to zero in a sigmoid fashion, with a longer lag and slower decrease for the nuclear compartment. Calculating the variance over multiple runs of the simulation shows that there is a peak in variance of cluster and monomer numbers between runs corresponding to a point a little after halfway through cluster depletion. (Figure 5.12).

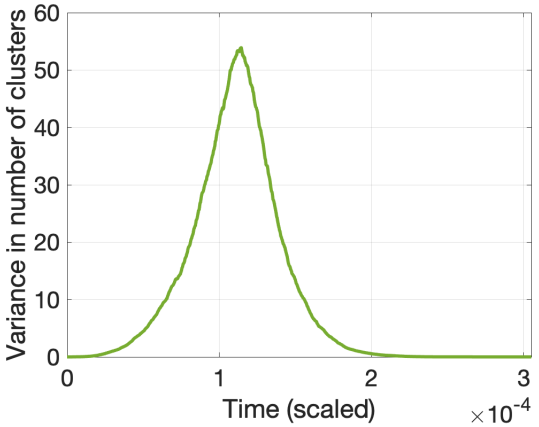
When considering the sizes of individual clusters over time, distinct phases of inflammasome formation and monomer depletion are evident (Figures 5.13 and 5.14). Initially, clusters cannot join since this is only possible via ASC-ASC interactions; there is therefore a lag in cluster growth while ASC monomers join to NLR oligomers, establishing the ASC filaments required for further ASC-ASC interactions. Once clusters have amassed sufficient ASC, the first joining event occurs (as in previous chapters, the first joining time is referred



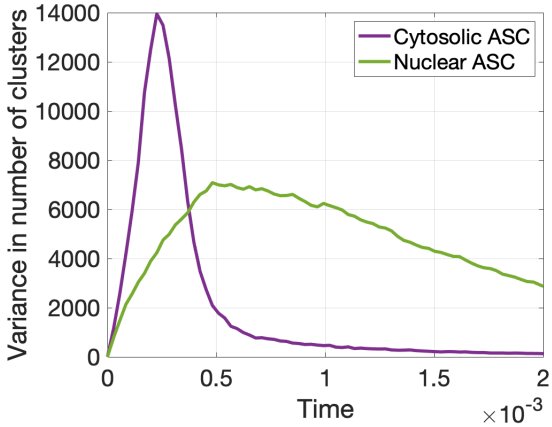
Number of clusters over time. Results for individual runs are plotted in black, and mean number of clusters is show in red.



Number of monomers over time. Results for individidual runs are shown in green for nuclear ASC and blue for cytosolic ASC.



Variance in number of clusters over time



Variance in number of monomers over time.

Figure 5.12: Output of 1000 simulations of the full inflammasome model with baseline parameter values as in Table 5.2. The mean number of clusters tends to 1, and the numbers of monomers in each compartment tends to 0; as for the ODE models, there is an initial lag in depletion. Variance between simulation runs increases until around halfway through ASC speck formation, then decreases to zero.

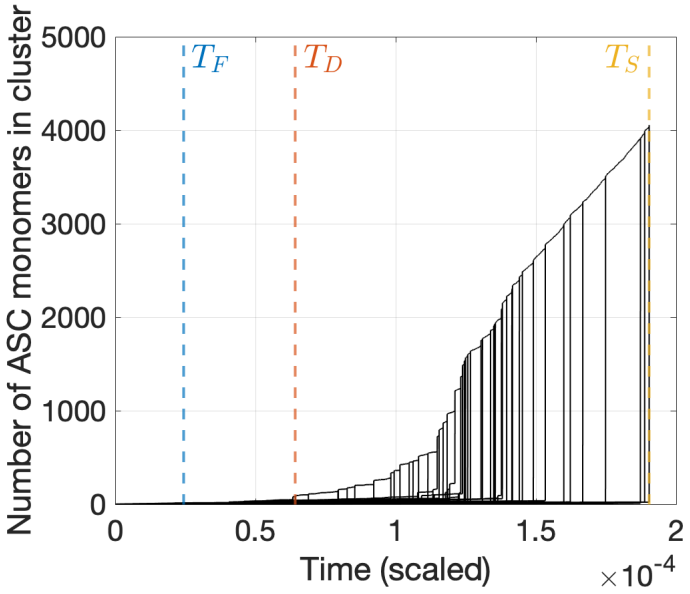
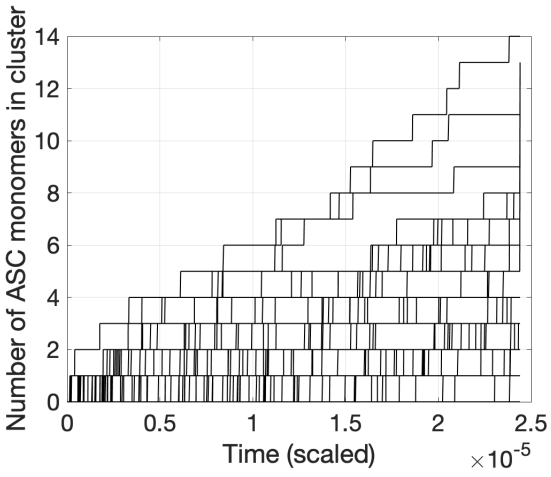


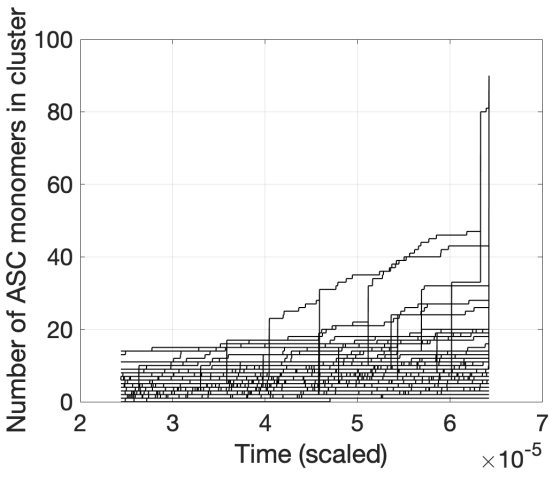
Figure 5.13: Size of individual clusters over time for one simulation run with baseline parameter values. The size of each cluster is plotted with a black solid line; vertical black lines indicate when two clusters join. The vertical dashed blue line indicates the first join time, and the vertical dashed red line indicates the dominant cluster emergence time. The single cluster emergence time is shown with a yellow dashed line. Initial growth is slow until after the dominant cluster emergence time, when there is a rapid increase in the size of one cluster.

to as T_F). A second phase then begins, in which clusters begin to join, although addition of ASC monomers also continues concurrently. Eventually, the emergence of a single dominant large cluster (at dominant cluster emergence time T_D) marks the beginning of a third stage. This cluster contains significantly more ASC monomers than any other cluster, and eventually absorbs all other, smaller clusters. The amalgamation of all NLR oligomers into one single, large cluster (at the single cluster emergence time T_S) marks the initiation of a final phase, in which no further joining events can occur and dynamics of the system are determined only by the depletion of ASC monomers as they are all eventually incorporated into the final inflammasome structure.

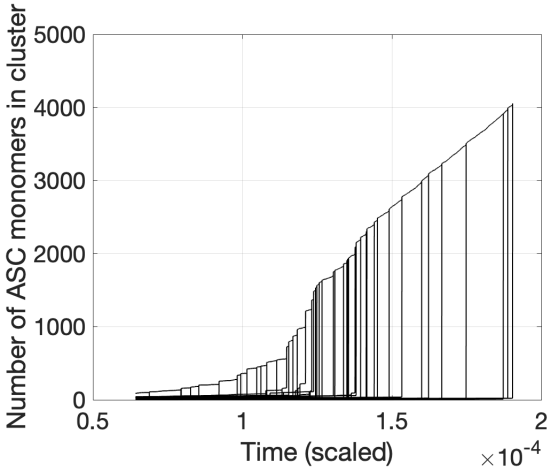
The initial growth of clusters relates closely to deterministic dynamics, as does growth following the emergence of a single cluster. However, the dominant cluster and single cluster emergence times are governed by stochastic dynamics arising during the second phase of cluster growth. This can be seen by observing that there is very little correlation between first joining time and dominant cluster or single cluster emergence time; however, there is a slight positive correlation between dominant cluster and single cluster emergence times (Tables 5.3 and 5.4). This suggests a large amount of variation in the length of the second phase of inflammasome growth. The variation in the length of the second phase leads



(a) Phase 1



(b) Phase 2



(c) Phase 3

Figure 5.14: Simulated run illustrated in Figure 5.13 divided into individual phases. Phase one shows gradual growth of all clusters. Phase 2 shows clusters beginning to join, but there is a fairly even distribution of cluster sizes. Phase 3 shows the emergence of one single large cluster; all other clusters are very small in comparison.

	$t_{1/10}$ cytosol	$t_{1/2}$ cytosol	$t_{1/10}$ nucleus	$t_{1/2}$ nucleus	T_F	T_D	T_S
$t_{1/10}$ cytosol	1	0.296	0.089	0.179	−0.011	0.066	0.062
$t_{1/2}$ cytosol	0.296	1	0.064	0.212	0.008	0.204	0.135
$t_{1/10}$ nucleus	0.089	0.064	1	0.420	0.035	0.008	0.009
$t_{1/2}$ nucleus	0.179	0.212	0.420	1	0.007	0.052	0.105
T_F	−0.012	0.008	0.035	0.007	1	0.087	0.044
T_D	0.066	0.204	0.008	0.052	0.087	1	0.070
T_S	0.0623	0.135	0.009	0.105	0.044	0.070	1

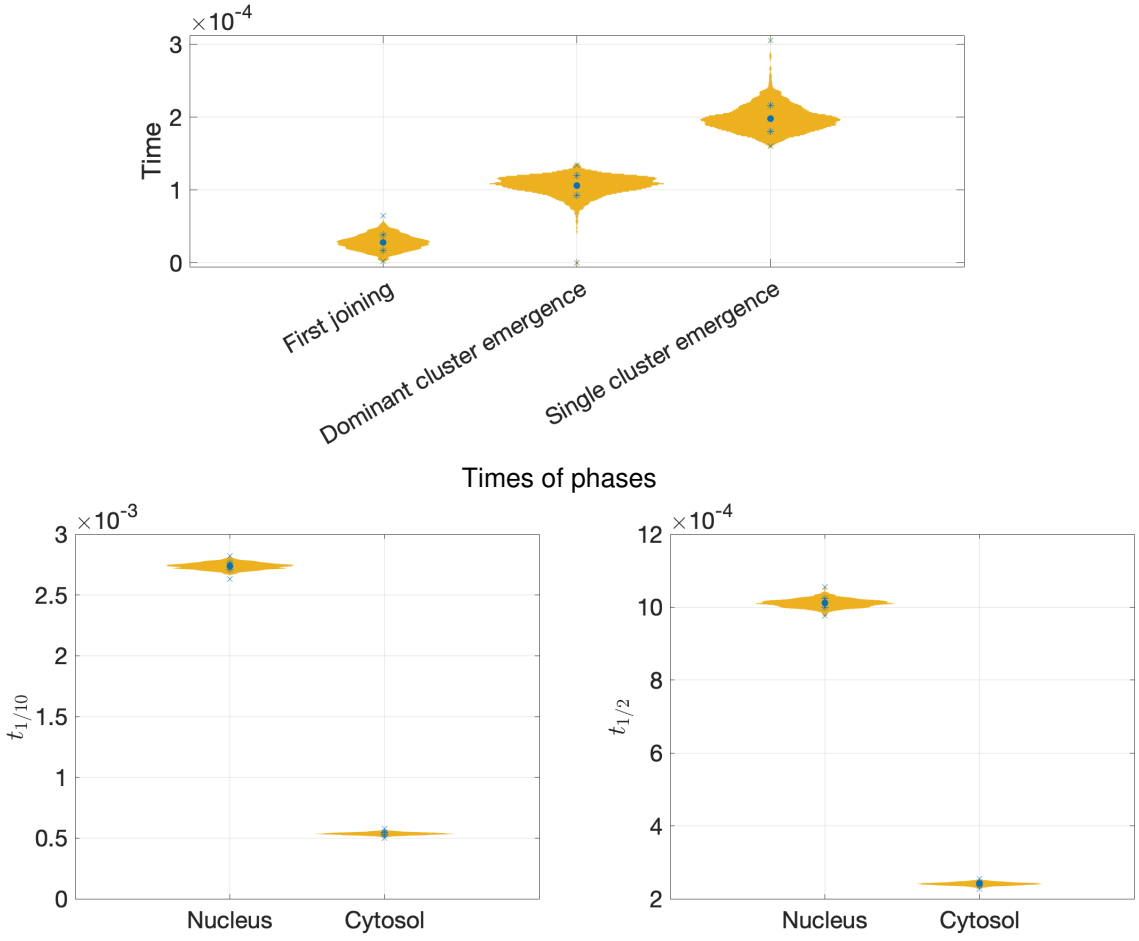
Table 5.3: ρ values for Spearman correlation coefficient between characteristic times for 1000 simulation runs with baseline parameters. There is a positive correlation between all monomer depletion characteristic times, as well as between first join time T_F, dominant cluster emergence time T_D and single cluster emergence time T_S.

to a considerable positive skew on the distribution of dominant cluster and single cluster emergence times, with a 'tail' of longer single cluster emergence times (Figure 5.15). This is observed in experimental data, and may account for the non-formation of inflammasomes by some cells, since for some cells inflammasome formation may take so long that it is interrupted by other cellular processes.

As with the experimental data, depletion of ASC monomers occurs more quickly for cytosolic than nuclear traces; using a one-sided Wilcoxon signed rank test, $p \approx 0$ for both t_{1_2} and $t_{1/10}$, suggesting that $t_{1/2}$ and $t_{1/10}$ are larger for nuclear traces. However, there is considerable correlation between all monomer depletion characteristic times for both the cytosolic and nuclear regions, suggesting that stochastic variation is less of a factor for monomer depletion dynamics than cluster growth dynamics.

	$t_{1/10}$ cytosol	$t_{1/2}$ cytosol	$t_{1/10}$ nucleus	$t_{1/2}$ nucleus	T_F	T_D	T_S
$t_{1/10}$ cytosol	1	$< 10^{-8}$	0.005	1.16×10^{-8}	0.738	0.037	0.049
$t_{1/2}$ cytosol	$< 10^{-8}$	1	0.042	$< 10^{-8}$	0.794	$< 10^{-8}$	1.72×10^{-5}
$t_{1/10}$ nucleus	0.005	0.042	1	$< 10^{-8}$	0.262	0.798	0.787
$t_{1/2}$ nucleus	1.15×10^{-8}	$< 10^{-8}$	$< 10^{-8}$	1	0.832	0.099	8.66×10^{-4}
T_F	0.738	0.794	0.262	0.832	1	0.006	0.163
T_D	0.037	$< 10^{-8}$	0.798	0.099	0.006	1	0.028
T_S	0.049	1.72×10^{-5}	0.787	8.66×10^{-4}	0.163	0.028	1

Table 5.4: p-values for Spearman correlation coefficient between characteristic times for 1000 simulation runs with baseline parameters. There is a significant correlation between all monomer depletion characteristic times, as well as between dominant cluster emergence time T_D and single cluster emergence time T_S, but not between the first joining time T_F and T_S, or between T_D and T_F.



$t_{1/10}$ for nuclear and cytosolic traces

$t_{1/2}$ for nuclear and cytosolic traces

Figure 5.15: Violin plots of distributions of characteristic times for 1000 simulation runs as described in Figure 5.12. Mean value is shown with a blue circle, mean pm standard deviation is shown with a blue star, and maximum and minimum values are shown with blue crosses. Single cluster emergence times show considerable positive skewness. Monomer depletion times for cytosol and nucleus are fairly symmetrically distributed; depletion of nuclear ASC is slower than depletion of cytosolic ASC.

5.3.3 Latin hypercube parameter sampling and characteristic time distributions

I will now consider the results for the Latin hypercube parameter sampling in order to gain a broader insight into the roles of the different parameters in the model. I will first outline the results for the full model as initially presented ('basic' assumptions), and then compare with simulations of variants with different initial conditions (i.e. models with no filament branching, one initial NLR oligomer, or random initial oligomer sizes).

Fitting a generalized linear regression model to basic simulation results

We can investigate the roles of the different parameters in the behaviour of characteristic times using regression techniques. Since by construction the Latin hypercube gives uncorrelated parameter values, we can investigate how each characteristic time varies over each parameter value. Standard linear regression is less useful here, due to the nonlinear relationships between parameters and characteristic times, and also non-normality of characteristic times (characteristic times cannot take negative values, and many parameter sets show positive skewness in the distribution of some characteristic times, especially single cluster emergence). However, a generalized linear model (GLM) can be used instead. This is not intended to be predictive; it need not be a perfect model for the simulation results, but fitting a model like this can at least give some insight into the role each parameter plays in determining results.

In this framework, for a characteristic time Y and set of parameters X, we have:

$$\mathbb{E}(Y|X) = g^{-1}(X\beta) \tag{5.3}$$

where g is an invertible 'link function' chosen prior to fitting, and $X\beta$ is the linear predictor of the data (β is a vector which specifies the linear relationship between the parameters and the mean value for a given set of parameters, and is determined through fitting the model).

A distribution for Y is also specified; in this case I have used the gamma distribution, which is often used for positive data with strong positive skewness. I have used a link function of $g(x) = -1/x$; this was the most effective at fitting the data from a range of link functions often used with the gamma distribution (compared with $g(x) = \log(x)$ and $g(x) = x$), and also has the advantage of being the 'canonical' link function for the gamma distribution.

I have also included values of pairwise products of parameters ($c_2 \times c_3, c_2 \times p, c_3 \times p$ etc.)

in the set of dependent variables in X to allow for the fact that there may be interactions between variables.

We can now fit GLMs to characteristic times extracted from our simulated data. We wish to derive a model which fits the data well while limiting the total number of parameters which play a significant role in the values of the characteristic times. Since there are six parameters, as well as 15 interactions resulting from pairwise combinations of those parameters, this gives a total of 2^{6+15} possible sets of dependent variables for the model. Instead of fitting each of these individually, I used a stepwise GLM fitting procedure (using the `stepwiseglm` function in Matlab) to choose an optimal model; this process starts with a minimal model, and systematically adds or removes dependent variables depending on the effect on the overall likelihood of the model. This process favours models with fewer dependent variables, while optimising the goodness of fit.

In particular, `stepwiseglm` starts from a model with only a single, constant term. At each step of the algorithm, the model is fitted, and the Akaike Information Criterion (AIC) is then used to determine whether terms should be added or removed from the model. The AIC is a measure of goodness of fit which penalises models for including increasing numbers of parameters to avoid overfitting; for a model with k parameters and maximum likelihood $\hat{L}$, the AIC is calculated as follows:

$$AIC = 2k - 2\log(\hat{L}) \tag{5.4}$$

In general, models with a lower AIC are preferred, since this may indicate a better fit and/or lower number of fitted parameters.

The algorithm used by the `stepwiseglm` function is as follows:

1. Fit initial model
2. Consider whether any terms not in the model should be added; if adding any terms leads to smaller AIC, add these to the model. Then proceed to Step 3.
3. Consider whether any terms in the model should be removed. If the removal of any terms leads to a smaller AIC, remove the term leading to the largest reduction and return to Step 2.

The algorithm terminates when the addition or removal of any parameter does not result in an improved AIC value.

	Estimate	Standard error	t-statistic	p-value
Intercept	43.155	8.453	5.105	3.379×10^{-7}
$c_3 \times m_1$	4.429	0.191	23.241	$< 10^{-8}$

Table 5.5: Coefficients in GLM for single cluster emergence times derived from Latin hypercube simulated data (100 simulations for 100 parameter sets); only the branching rate c_3 and the rate of movement of ASC out of the nucleus m_1 contribute significantly. Error degrees of freedom: 7729.

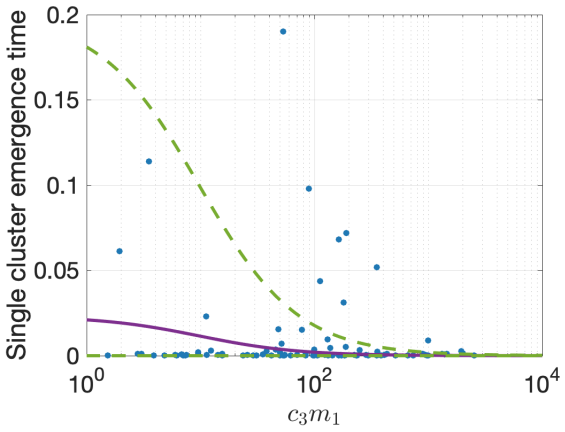


Figure 5.16: Illustration of GLM fit for single cluster emergence time to Latin hypercube simulations (100 repeat simulations for 100 parameter sets). Mean single cluster emergence times for each parameter set are plotted against the value of $c_3 m_1$ (the product of branching rate and rate of movement of ASC out of the nucleus); these are shown as blue points. The mean value predicted by the GLM is shown as a solid purple line, and a 95% confidence interval as predicted by the GLM lies between the dotted green lines. The mean and variance of the fitted gamma distribution decrease as $c_3 X_0$ increases.

Inflammasome formation times

In the case of the single cluster emergence time T_S, the resultant model has $R^2 = 0.0478$ and only has two terms: a constant term, and a term describing an interaction between the value of c_3 and m_1. A description of the coefficients is given in Table 5.5. The dispersion parameter is $\phi = 6.939$; this is defined as $\phi := 1/\alpha$, where α is the shape parameter of the gamma distribution.

This suggests that the best-fitting model is as follows:

$$\mathbb{E}(T_S) = \frac{1}{43.155 + 4.429 c_3 m_1} \tag{5.5a}$$

$$T_S \sim \Gamma\left(\frac{1}{6.939}, \frac{43.155 + 4.429 c_3 m_1}{6.939}\right) \tag{5.5b}$$

Note that this does not suggest that this is an exact expression for the distribution of single cluster emergence time, but only that this is the most appropriate model of this form. This

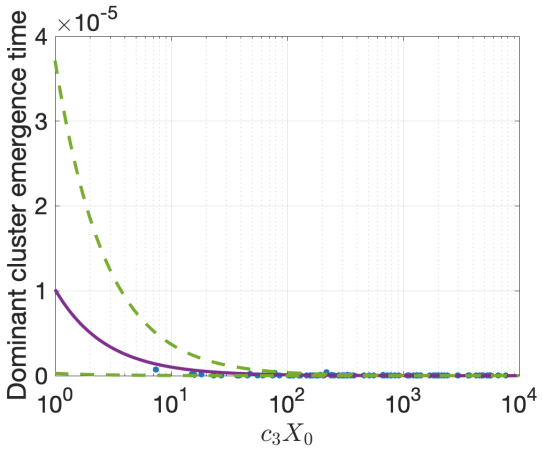


Figure 5.17: Illustration of GLM fit for dominant cluster emergence time for simulated Latin hypercube data (100 simulation runs for 100 parameter sets). Mean single cluster emergence times for each parameter set are plotted against the value of c_3X_0 (the product of branching rate and the initial number of clusters); these are shown as blue points. The mean value predicted by the GLM is shown as a solid purple line, and a 95% confidence interval as predicted by the GLM lies between the dotted green lines. The mean and variance of the fitted gamma distribution decrease as c_3X_0 increases.

suggests that the most important parameters in determining the emergence times of a single cluster are c_3 and m_1, and that the mean emergence time decreases as c_3 and m_1 increase (Figure 5.16). The shape of the distributions of data also changes with the parameter values; as c_3 and m_1 increase, the scale parameter β increases, leading to reduced variance of results. The large-time 'tail' of the distribution also decreases.

This result is in contrast to the results for single cluster emergence time in the simplified model, presented in Section 4.3.1; in particular, this analysis demonstrated that the cluster joining rate c_2 was the main factor in determining the distribution of the single cluster emergence time in the simplified model. This difference arises from the fact that the cluster joining rate in the simplified model is independent of cluster size, whereas in the full model, cluster joining is dependent on the number of filaments in the relevant clusters.

The model fit is somewhat different for the dominant cluster emergence time T_D; in this case we have $R^2 = 0.4249$ and $\phi = 0.9786$. Details of the coefficients in these models are given in Table 5.6. In comparison to the single cluster emergence times, the only significant term identified is an interaction between c_3 and X_0. This suggests that the mean and variance of the distribution of dominant cluster emergence time decreases as c_3 and X_0 increase, but that m_1 is less likely to be important. Thus, the larger the initial number of NLR oligomers, the shorter the time for a dominant cluster to emerge.

Together, these results suggest that X_0 is important for determining dominant cluster

	Estimate	**Standard error**	t**-statistic**	p**-value**
Intercept	106488.488	109270.417	0.975	0.330
$c_3 \times X_0$	98428.245	1105.533	89.032	$< 10^{-8}$

Table 5.6: Coefficients in GLM for dominant cluster emergence times derived from Latin hypercube simulated data (100 simulations for 100 parameter sets). Only the branching rate c_3 and the initial number of clusters X_0 contribute significantly. Error degrees of freedom: 9898.

emergence time, but less so for single cluster emergence time, when the availability of free monomers (which is controlled by m_1 at late times) becomes much more important. However, it is also worth bearing in mind that this model may not be a very precise fit for the simulated data (see for instance the fairly low values of R^2), suggesting that the conclusions we draw here may be limited.

Meanwhile, the results for first joining time are different again. For this model, $R^2 = 0.8209$ and $\phi = 0.1596$, and details of the coefficients are given in Table 5.7. The model contains all parameters except p, as well as all interactions between those variables (with the exception of $c_3 \times X_0$). It is unsurprising that c_2, the parameter which controls joining rate, is so important, as is the initial number of clusters X_0; the same was observed for the first joining time in the simplified model. However, in the full model, the branching rate and the parameters controlling movement of ASC between compartments also influence the distribution of the first joining time. It is also notable that the dispersion parameter is much lower than for the fitted models for single and dominant cluster emergence times; since the skewness of a gamma distribtion is given by $2\sqrt{\alpha} = 2 \times \sqrt{1/\phi}$, this suggests that the skewness of the distribution of first joining time is much smaller than the skewness of the distributions of the other characteristic times examined previously. This further supports the idea that the skewness that arises for statistics describing inflammasome formation times arise from late-time dynamics.

Depletion of ASC monomers

We now turn our attention to the monomer depletion characteristic times; in particular, $t_{1/2}$ and $t_{1/10}$ for the nuclear and cytosolic regions.

The fitted GLMs for $t_{1/2}$ and $t_{1,/10}$ have $R^2 = 0.9955, \phi = 0.0063$ and $R^2 = 0.9963, \phi = 0.0027$ respectively for the cytosolic region, and $R^2 = 0.9951, \phi = 0.0072$ and $R^2 = 0.9501, \phi =$

	Estimate	**Standard error**	t**-statistic**	p**-value**
Intercept	-2205.614	470.652	-4.686	2.819×10^{-6}
c_2	1374.331	378.581	3.630	2.846×10^{-4}
c_3	7268.171	420.389	17.289	$< 10^{-8}$
m_1	34.072	3.338	10.208	$< 10^{-8}$
m_1/m_2	2145.964	570.788	3.760	1.711×10^{-4}
X_0	109.589	2.235	49.041	$< 10^{-8}$
$c_2 \times c_3$	-260.598	86.869	-3.000	0.003
$c_2 \times m_1$	-1.829	0.938	-1.950	0.051
$c_2 \times m_2/m_1$	1331.628	289.235	4.604	4.196×10^{-6}
$c_2 \times X_0$	48.134	0.930	51.737	$< 10^{-8}$
$c_3 \times m_1$	-3.006	0.709	-4.241	2.240×10^{-5}
$c_3 \times m_2/m_1$	-2240.843	277.869	-8.064	$< 10^{-8}$
$m_1 \times m_2/m_1$	7.794	2.197	3.548	3.900×10^{-4}
$m_1 \times X_0$	-0.086	0.005	-17.371	$< 10^{-8}$
$m_1 \times X_0$	-5.000	1.716	-2.914	0.004

Table 5.7: Coefficients in GLM for first joining times derived from Latin hypercube simulated data (100 simulations for 100 parameter sets). All parameters except the branching rate p are significant in determining the first joining time. Error degrees of freedom: 9985.

0.048 respectively for the nuclear region. Note the large values of R^2, which suggest a particularly good model fit. Details of the coefficients are given in Tables 5.8, 5.9, 5.10 and 5.11. All models contain all six parameters, as well as most pairwise interactions between parameters. This is somewhat in contrast to the simplified model, in which the cluster joining rate c_2 does not affect the monomer depletion characteristic times, and naturally the rates of movement between cellular compartments were not included in this model. This shows that the movement of ASC monomers between cellular compartments are significant in determining the dynamics of ASC monomer depletion, supporting the inclusion of the two spatial compartments in the model.

Having established the key differences arising from different parametrisations with basic modelling assumptions, we will now consider comparisons to alternative models. In Chapter 6, all of the simulation results described above will be compared to experimental data in order to ascertain which parametrisations give rise to the closest match.

5.3.4 Comparison with alternative models

We now consider the same set of parameters from our Latin hypercube sampling, and use them to run the full model with various modifications. In all cases, we can compare the distributions of each characteristic time for each parameter set; in particular, we can use the Wilcoxon rank-sum test, a nonparametric test with the null hypothesis that two sets of data have the same median, and the Kolmogorov-Smirnov test, a nonparametric test with the null hypothesis that two sets of data are drawn from the same distribution. Before performing Kolmogorov-Smirnov tests, I have subtracted the sample means for each parameter grouping in order to remove dependence of the results on the centralisation of data.

I have chosen to use nonparametric tests since, as we have already established, characteristic times often have clearly non-normal distributions with considerable positive skewness and different variances.

These tests allow us to identify significant differences between distributions, and pinpoint the differences arising from each model modification.

Removing branching from the model

We will first consider a set of simulations in which all parameters are as in the simulations for the basic assumptions, but with $c_3 = 0$ in all cases; this corresponds to removing branch-

	Estimate	**Standard error**	t**-statistic**	p**-value**
Intercept	-95.138	18.842	-5.049	4.513×10^{-7}
c_2	-51.918	2.519	-20.611	$< 10^{-8}$
c_3	2249.963	18.764	119.910	$< 10^{-8}$
p	187.772	8.766	21.420	$< 10^{-8}$
m_1	0.170	0.092	1.843	0.065
m_1/m_2	-89.878	21.793	-4.124	3.751×10^{-5}
X_0	12.606	0.059	212.015	$< 10^{-8}$
$c_2 \times c_3$	44.868	2.819	15.919	$< 10^{-8}$
$c_2 \times p$	43.179	1.222	35.344	$< 10^{-8}$
$c_2 \times m_1$	0.078	0.010	8.047	$< 10^{-8}$
$c_2 \times X_0$	-0.172	0.014	-12.553	$< 10^{-8}$
$c_3 \times p$	-144.048	3.166	-45.511	$< 10^{-8}$
$c_3 \times m_1$	-0.113	0.030	-3.761	0.0001
$c_3 \times m_2/m_1$	-103.850	10.577	-9.818	$< 10^{-8}$
$c_3 \times X_0$	1.660	0.024	70.156	$< 10^{-8}$
$p \times m_1$	0.264	0.020	12.992	$< 10^{-8}$
$p \times m_2/m_1$	178.124	6.718	26.515	$< 10^{-8}$
$p \times X_0$	2.097	0.016	132.247	$< 10^{-8}$
$m_1 \times m_2/m_1$	0.241	0.060	4.010	6.112×10^{-5}
$m_1 \times X_0$	-0.003	1.328×10^{-4}	-21.817	$< 10^{-8}$
$m_2/m_1 \times X_0$	-0.320	0.050	-6.449	$< 10^{-8}$

Table 5.8: Coefficients in GLM for monomer depletion characteristic time $t_{1/2}$ in the cytosolic region derived from Latin hypercube simulated data (100 simulations for 100 parameter sets); all parameters are significant in determining $t_{1/2}$. Error degrees of freedom: 9979.

	Estimate	**Standard error**	t**-statistic**	p**-value**
Intercept	138.959	41.340	3.361	7.785×10^{-4}
c_2	-437.162	23.512	-18.593	$< 10^{-8}$
c_3	3173.348	33.949	93.473	0
p	273.207	15.693	17.410	$< 10^{-8}$
m_1	-4.321	0.196	-22.029	$< 10^{-8}$
m_1/m_2	-797.637	51.486	-15.492	$< 10^{-8}$
X_0	93.890	0.198	474.001	0
$c_2 \times c_3$	131.526	4.993	26.344	$< 10^{-8}$
$c_2 \times p$	120.931	3.263	37.062	$< 10^{-8}$
$c_2 \times m_1$	0.256	0.035	7.320	$< 10^{-8}$
$c_2 \times m_2/m_1$	250.453	16.231	15.431	$< 10^{-8}$
$c_2 \times X_0$	-0.248	0.044	-5.662	1.538×10^{-8}
$c_3 \times p$	-163.143	6.536	-24.962	$< 10^{-8}$
$c_3 \times m_1$	0.279	0.063	4.450	8.680×10^{-6}
$c_3 \times m_2/m_1$	-522.830	20.951	-24.955	$< 10^{-8}$
$c_3 \times X_0$	2.491	0.047	53.158	$< 10^{-8}$
$p \times m_1$	0.333	0.048	7.000	$< 10^{-8}$
$p \times m_2/m_1$	804.867	15.218	52.889	$< 10^{-8}$
$p \times X_0$	3.030	0.041	73.655	$< 10^{-8}$
$m_1 \times m_2/m_1$	2.202	0.154	14.340	$< 10^{-8}$
$m_2/m_1 \times X_0$	-1.528	0.149	-10.261	$< 10^{-8}$

Table 5.9: Coefficients in GLM for monomer depletion characteristic time $t_{1/10}$ in the cytosolic region derived from Latin hypercube simulated data (100 simulations for 100 parameter sets); all parameters are significant in determining $t_{1/10}$. Error degrees of freedom: 9979.

	Estimate	**Standard error**	t**-statistic**	p**-value**
Intercept	2.548	0.264	9.654	$< 10^{-8}$
c_2	0.522	0.045	11.537	$< 10^{-8}$
c_3	-0.235	0.050	-4.738	2.191×10^{-6}
p	0.614	0.085	7.242	$< 10^{-8}$
m_1	1.192	0.004	308.794	$< 10^{-8}$
m_1/m_2	0.219	0.205	1.069	0.285
X_0	-6.612×10^{-4}	3.487×10^{-4}	-1.896	0.058
$c_2 \times c_3$	-0.028	0.006	-4.896	9.922×10^{-7}
$c_2 \times p$	-0.072	0.009	-8.369	$< 10^{-8}$
$c_2 \times m_1$	-0.003	7.104×10^{-4}	-3.984	6.825×10^{-5}
$c_2 \times X_0$	-5.338×10^{-4}	6.847×10^{-5}	-7.797	$< 10^{-8}$
$c_3 \times p$	-0.129	0.010	-12.569	5.880×10^{-36}
$c_3 \times m_1$	0.014	6.560×10^{-4}	21.671	$< 10^{-8}$
$c_3 \times m_2/m_1$	0.214	0.032	6.690	$< 10^{-8}$
$p \times m_1$	0.012	6.725×10^{-4}	17.824	$< 10^{-8}$
$p \times X_0$	-7.235×10^{-4}	8.580×10^{-5}	-8.537	$< 10^{-8}$
$m_1\ timesX_0$	$1/327 \times 10^{-4}$	5.792×10^{-6}	22.916	$< 10^{-8}$
$m_2/m_1 \times X_0$	-6.873×10^{-4}	3.214×10^{-4}	-2.138	0.033

Table 5.10: Coefficients in GLM for $t_{1/2}$ in the nuclear region derived from Latin hypercube simulated data (100 simulations for 100 parameter sets); all parameters are significant in determining $t_{1/2}$. Error degrees of freedom: 9982.

	Estimate	**Standard error**	t**-statistic**	p**-value**
Intercept	18.019	4.216	4.274	1.940×10^{-5}
c_2	9.553	0.940	10.167	$< 10^{-8}$
c_3	1.620	1.527	1.061	0.289
p	15.829	1.428	11.088	$< 10^{-8}$
m_1	4.421	0.049	89.532	$< 10^{-8}$
m_1/m_2	26.015	3.856	6.747	$< 10^{-8}$
X_0	0.025	0.006	4.475	7.719×10^{-6}
$c_2 \times c_3$	-0.753	0.111	-6.767	$< 10^{-8}$
$c_2 \times p$	-1.283	0.138	-9.263	$< 10^{-8}$
$c_2 \times m_1$	-0.033	0.009	-3.546	3.923×10^{-4}
$c_2 \times m_2/m_1$	-0.559	0.528	-1.058	0.290
$c_2 \times X_0$	-0.010	0.001	-8.350	$< 10^{-8}$
$c_3 \times p$	-2.747	0.162	-16.988	$< 10^{-8}$
$c_3 \times m_1$	0.351	0.010	33.834	$< 10^{-8}$
$c_3 \times m_2/m_1$	1.681	0.667	2.518	0.012
$c_3 \times X_0$	-0.006	0.001	-4.647	3.405×10^{-6}
$p \times m_1$	0.189	0.008	19.438	$< 10^{-8}$
$p \times m_2/m_1$	-1.050	0.554	-1.895	0.058
$p \times X_0$	-0.017	0.001	-12.709	$< 10^{-8}$
$m_1 \times X_0$	0.002	$8.053e \times 10^{-5}$	30.135	$< 10^{-8}$
$m_2/m_1 \times X_0$	-0.023	0.006	-3.736	1.879×10^{-4}

Table 5.11: Coefficients in GLM for $t_{1/10}$ in the nuclear region derived from Latin hypercube simulated data (100 simulations for 100 parameter sets); all parameters are significant in determining $t_{1/10}$. Error degrees of freedom: 9979.

ing reactions from the model. One-sided Wilcoxon rank sum tests show that all characteristic times with the exception of the dominant cluster emergence times are significantly lower for the basic assumptions compared to the simulations without filament branching ($p < 0.05$), although this may be due to the fact that setting $c_3 = 0$ decreases the overall monomer depletion and cluster growth rates, since the filament extension rate p is unchanged. However, for some parametrisations there is a significant difference between the shapes of the distributions of single cluster emergence times for the basic assumptions simulations and the simulations without branching, even after the distributions have been centred by mean (Kolmogorov-Smirnov test, $p < 0.05$). There is a strong correlation between the value of c_3 in the basic assumption simulations and the Kolmogorov-Smirnov test statistic ($\rho = 0.615, p < 10^{-8}$); unsurprisingly, this suggests that the largest differences in distribution arise when c_3 is large in the basic assumptions dataset.

We also see differences in distribution between the basic assumptions simulations and simulations without joining for $t_{1/2}$ in the cytosolic region and $t_{1/10}$ in the nuclear region; this suggests that branching may be important in determining the levels of ASC in the nuclear region at early times, but has a larger effect on monomer depletion in the cytosolic regions at later times. In particular, there is a strong correlation between c_3 and the Kolmogorov-Smirnov statistic when comparing distributions of $t_{1/2}$ in the cytosolic region ($\rho = 0.659, p < 10^{-8}$ and $t_{1/10}$ in the nuclear region ($\rho = 0.362, p = 2.16 \times 10^{-4}$). This is due to the fact that, as observed previously, as c_3 increases the spread of the distribution of single cluster emergence time decreases.

There is thus clearly a significant difference in the distributions of characteristic times between the models with and without branching. The simulations without branching will be compared to experimental data alongside the basic assumptions simulations in the following chapter.

Removing joining from the model

We will now consider a modification in which we assume there is a single initial NLR oligomer; we assume that this contains 10 monomer subunits. Naturally, this removes joining from the model (if we set $c_2 = 0$ we could still start with multiple oligomers, but there would be no possible mechanism by which a single inflammasome could emerge).

We now fix $X_0 = 1$ and otherwise use the Latin hypercube sample parameter sets,

keeping the total number of NLRs the same as in the basic model. We will compare only the characteristic statistics related to ASC monomer counts, i.e. $t_{1/2}$ and $t_{1/10}$ in both the nuclear and cytosolic compartments (the other characteristic times are no longer relevant to this model).

In virtually all cases, comparing the distribution of characteristic times between the full model and one-oligomer model gives a significant difference in the medians of $t_{1/2}$ and $t_{1/10}$ in the nuclear compartment, and in $t_{1/2}$ in the cytosolic compartment. However, for $t_{1/10}$ in the cytosolic compartment, there is no significant difference in median for the vast majority of parameter sets. Using right-tailed Wilcoxon rank sum tests shows that for the majority of the parameter sets, removing joining from the model leads to an increase in the median value of $t_{1/2}$ or $t_{1/10}$ in the nuclear region, and for $t_{1/2}$ in the cytosolic region. There is no significant difference in the shapes of distribution for $t_{1/2}$ and $t_{1/10}$ in the cytosolic compartment, and around half of the parameter sets show a significant difference in distribution shape for $t_{1/2}$ and $t_{1/10}$ in the nuclear compartment.

This suggests that the inclusion of joining into the system does not significantly affect the shape of the distribution of characteristic times in the cytosolic region, but that the median values are changed; the overall rate of monomer depletion is decreased when joining is removed. However, it is likely that the decrease in the total number of NLR oligomers is responsible for these changes. The fact that this set of simulations leads to results very similar to the basic assumption but shifted by mean suggests that the case where exactly one NLR oligomer is initially present may not be distinguishable from other initial numbers of NLR oligomers. Nonetheless, I will include this dataset in the comparisons with data in the following chapter.

Random initial conditions

Considering now the dataset in which we have random initial conditions, we can compare the results for each parameter set with those for the basic model.

For most parameter sets, emergence times of a single cluster have a different median (Wilcoxon rank test, $p < 0.05$), while the distribution of other characteristic times is broadly unaffected. Likewise, after centering the distributions for both datasets by subtracting the median, of all characteristic times only the shape of the distribution of emergence times of a single cluster differs significantly from the corresponding distribution in the full model without

modification (Kolmogorov-Smirnov test, $p < 0.05$). This suggests that changing the sizes of the initial oligomers does not have a significant effect on any characteristic time except for emergence time of a single cluster.

Comparing now the Kolmogorov-Smirnov test statistic for each parameter set with the size of those parameters, a dependence emerges between the value of the Wilcoxon rank test statistic and of the parameter X_0 ($\rho = -0.857, p = 2.9326 \times 10^{-21}$). This suggests that the lower the value of X_0, the higher the value of the test statistic, and thus the larger the difference in median between the full model and the modification with random ICs. Importantly, the total number of NLRs is the same in both models, and it is only the division of NLRs between oligomers which is random; this suggests that the smaller the initial number of oligomers, the larger the effect allowing oligomers of different sizes may have on the results compared to assuming all oligomers are the same size. Likewise, for the Kolmogorov-Smirnov test comparing the shape of the distribution of single cluster emergence time, only X_0 shows a significant correlation with values of the test statistic ($\rho = 0.742, p = 2.0026 \times 10^{-13}, \tau = 0.553, p = 1.4844 \times 10^{-11}$). This suggests that the larger the value of X_0, the larger the difference in distribution between the full model without modification and the full model with random ICs; when there are more NLR oligomers initially, there is the potential for a much more uneven spread of monomers between oligomers, and thus in these cases it may take much longer for a single cluster to emerge; this will lead to a longer 'tail' for large times in the distribution of single cluster emergence time.

I will compare the results of the random ICs simulations to experimental data in Chapter 6.

5.4 Discussion

5.4.1 Summary

In this chapter, I have presented results for both the deterministic and stochastic formulations of the full model of inflammasome formation.

Overall, while the full deterministic model displays some similar behaviour to the simplified model, there are evidently more complex dynamics at play; in particular, the distinction between C1 and C3 processes (ASC monomers joining to free NLRs and ASC filament branching respectively) leads to clear phases in cluster growth and merging when the joining

rate of ASC to NLRs is slow. In particular, we see a clear initial 'lag' phase, in which growth and merging is limited due to a bottleneck in filament seeding from NLRs. I have defined a set of parameters for which the lag phase is not present, as in the simplified inflammasome model.

However, the full deterministic model still fails to capture some of the behaviour of inflammasome formation systems, especially at late times (the tendency of X to become negative is particularly undesirable). The full deterministic model still does not address the fundamental inability of ODE models to capture the behaviour of the inflammasome systems at late times when there are few clusters, at which point we would expect the cluster sizes to level off, and the variances of cluster sizes to tend towards zero.

Turning our attention to the stochastic formulation, there are a number of conclusions we can draw from the results of our simulations. By considering a large cohort of simulations with a basic parametrisation, we have seen that inflammasome formation naturally falls into several phases: an early phase, in which individual oligomers recruit ASC monomers, a second phase in which oligomers begin to join but the distribution of oligomer sizes remains fairly even, and a later phase in which a single dominant cluster emerges. Analysing correlations between the timings of these phases demonstrates that variation between cells in the same sample appears to arise because of the stochasticity caused by joining events in the second phase.

I have also run simulations of the model with a large range of parameter values and fitted a GLM to characteristic times for each parameter set, in order to determine which parameters are most important in determining each characteristic time. We generally see that with basic assumptions, increasing parameter values has a scaling effect, which in turn increases the rapidity of the inflammasome formation process as a whole; both the means and variances of characteristic times decrease as branching rate, filament extension rate and initial number of NLR oligomers increases. A particularly notable result is that characteristic times associated with earlier times in the inflammasome formation process (such as the first joining times and monomer depletion characteristic times) are significantly dependent on many parameters, whereas the characteristic times associated with later stages (single cluster and dominant cluster emergence) are dependent on few parameters. In particular, we see that the filament branching rate c_3 and the rate of movement of ASC monomers out of the nucleus m_1 are most important for single cluster emergence times, whereas c_3 and the initial number

of oligomers X_0 are more important for dominant cluster emergence time. This may give some insight into why multiple inflammasomes may sometimes be observed in systems with an overabundance of active NLRs; in particular, since the time of single cluster emergence is not significantly affected by X_0, but the time of emergence of a dominant cluster decreases as X_0 increases, this suggests that as X_0 increases the time interval between emergence of a dominant cluster and the final emergence of a single cluster also increases. It is then possible that other biological processes might be intervening in inflammasome formation before the dominant cluster emergence time, if this occurs particularly late. This may explain why systems which may have smaller initial numbers of NLRs present are less likely to display inflammasome formation in all cells (for example, in the NLRP3 inflammasome system).

I have also simulated results for the full model when slightly different assumptions have been used. In particular, I have simulated results for random initial NLR oligomer sizes (compared to the basic assumption, in which all NLR oligomers contain exactly 10 monomers), a system in which branching does not occur, and also a system in which only a single NLR oligomer is present. The simulations without branching and random oligomer sizes in particular are clearly distinct from the simulations for the basic assumptions. All four sets of simulations will be compared to experimental data in the following chapter, to determine which model gives rise to results closest to the experimental data.

5.4.2 Further work

Now that we have established the key differences between results for the simplified model and the full model, as well as the differences arising for different parametrisations of each model, the natural next step is to compare all of these results to the experimental data presented in Chapter 2. In particular, we can compare the distributions of characteristic times both to the distributions derived analytically for the stochastic simplified model in Chapter 4, and to the empirical distributions arising from the simulations in this chapter. This will be the subject of the following chapter.

One aspect of inflammasome formation which has not been particularly interrogated here is the effect of ASC overexpression; in particular, in some ASC overexpression systems we observe multiple ASC specks forming. This could be examined by running the simulations with a larger initial ASC count; however, this would be a time-consuming exercise since, in general, simulations of ASC overexpression systems take much longer to run, since more

reactions are inovlved. Thus, running simulations covering all of the parameter sets obtained from Latin hypercube sampling would be extremely slow. In order to make this feasible, it may be necessary to adapt the (fairly simple) simulation algorithm used here in order for the algorithm to handle the reactions more efficiently.

Chapter 6

Comparison of models of the inflammasome to experimental data

6.1 Introduction

So far, I have mainly considered the experimental dataset presented in Chapter 2, and the models of inflammasome formation presented in Chapter 3, in isolation; the next stage in the cycle of model development is to compare the models to experimental data directly. This allows us to assess whether the models accurately reflect the inflammasome formation process and identify areas for improvement; we will also be able to discriminate between the different variants of the models produced in order to ascertain which elements should be included in a cohesive conceptual model of inflammasome formation. In particular, I have defined characteristic times which summarise key features of the experimental data and the outputs from both the simplified and the full inflammasome models.

A key observation established in Chapters 4 and 5 is that, as ASC monomers are depleted, for some parametrisations the full model demonstrates an initial 'lag phase' before the maximum rate of decrease is reached, while the simplified model exhibits a sharp onset of monomer depletion without such a lag phase. In order to assess which of these models gives a better reflection of the experimental data, I fitted the traces describing ASC monomer depletion derived from the experimental data to functions approximating either gradual or steep depletion onsets at early times. I will demonstrate that the most successful fit is obtained with an initial sharp decrease onset, as in the simplified model; this suggests a minimal role for the initial formation of filaments in limiting inflammasome formation, and that initial ASC

recruitment rate by NLRs is comparable in magnitude to other cluster growth parameters.

I then fitted the characteristic times derived from the experimental dataset to the theoretical distributions derived from the results for the simplified model in Chapter 4; the best fits are given by models with a low number of NLR oligomers initially present, and with ASC filament branching included. This suggests that branching does indeed play a role in inflammasome formation, contrary to many of the predominant conceptual models.

Finally, I compared the characteristic time distributions derived from experimental data to simulations of the full model presented in Chapter 5; the closest fits arise from inflammasome systems with random initial sizes of NLR oligomers, ASC filament branching, and joining of clusters.

6.2 Ascertaining the importance of filament-driven processes from experimental data

A clear difference between the simplified and full models established in Sections 4.2.2 and 5.2 was that the simplified model shows an immediate sharp decrease in ASC monomer abundance following infection, whereas the full model can show an initial 'lag' before a steep decrease for small values of c_1. This lag arises because there is an initial time period in which filaments must be established before filament branching and formation can occur. The lag is not present in the simplified model since there is no initial filament formation stage, as NLRs are not treated as distinct from ASC filaments.

It is important to establish which model better captures this initial stage of inflammasome formation for several reasons. Firstly, the lag phase (if it is indeed present) may be biologically important, for example in introducing a delay before inflammasome formation begins to proceed in earnest, so that inflammasome formation could be reversed at very early times if necessary. Furthermore, if we can define the nature of a lag phase at the start of monomer depletion, this will allow us to obtain an estimate for the time between the introduction of the inflammasome trigger and the initiation of ASC monomer depletion. This is the starting point of the mathematical models analysed in previous chapters; in order to be able to compare the experimental data to the models, the time before ASC monomer depletion begins for each cell must therefore be subtracted from the other characteristic times prior to analysis.

Previously, in Chapter 2, I described two datasets (referred to as the FLICA and Mito-

Tracker datasets) consisting of time series imaging of cells in which ASC is fluorescently labelled undergoing inflammasome formation. A proxy for the total ASC monomer abundance in the nuclear and cytosolic regions can be obtained by integrating the intensity over these regions in each frame, resulting in intensity 'traces', which give the relative ASC abundance over time in each cell. In Section 2.3, I presented a trilinear fitting methodology to obtain intial estimates for the onset time of ASC monomer depletion. For each intensity measurement x_i in a trace we have:

$$x_i = \beta_0 + \beta_1 t_i + \beta_2 (t_i - t_{b_1})^+ + \beta_3 (t_i - t_{b_2})^+ + \epsilon_i \qquad \text{(2.1 revisited)}$$

where t_i is the time at which measurement x_i is taken, t_{b_1} and t_{b_2} are the start and end point of the interval of speck formation, β_0 is the initial intensity, β_1, β_2 and β_3 are the gradients of each section of the trilinear fit, and ϵ_i is the error at datapoint i.

$(t_i - t_{b_j})^+$ is defined as follows:

$$(t_i - t_{b_j})^+ = \begin{cases} t_i - t_{b_j} & \text{if } t_i > t_{b_j} \\ 0 & \text{otherwise} \end{cases} \qquad \text{(2.2 revisited)}$$

In the analysis in Section 2.3, t_{b_1} was taken as the time of ASC monomer depletion onset, with inflammasome formation taking place between t_{b_1} and t_{b_2}. The values of β_0, β_1, β_2 and β_3 were fitted for each possible pair of breaktimes t_{b_1} and t_{b_2} using maximum likelihood estimation. The pair of breaktimes t_{b_1} and t_{b_2} giving the largest overall maximum likelihood were then chosen as estimators for these breakpoints. Identifying the period of time over which ASC is decreasing was useful for comparing traces in the dataset, and for determining the rates of monomer depletion, as well as identifying traces which were too noisy to be analysed, and which could then be removed. However, now that we have ascertained that the onset may be sigmoid instead of exponential, it is worth refining this estimate of ASC monomer depletion onset time.

In order to ascertain whether a sigmoid or exponential fit at the onset of monomer depletion fits the data better, for each trace (both nuclear and cytosolic) in both the FLICA and MitoTracker datasets which met the trilinear fitting conditions laid out in Section 2.3, I identified the first breakpoint t_{b_1} from the initial trilinear fit. I took as my interval of interest the n datapoints to either side of t_{b_1}. I then fitted three different models to the datapoints in this

interval, as follows:

$$x_i = \beta_0 + \beta_1 t_i + \beta_2 (t_i - t_b)^+ + \epsilon_i \tag{6.1a}$$
$$x_i = \beta_0 + \beta_1 t_i + \beta_2 ((t_i - t_b)^+)^2 + \epsilon_i \tag{6.1b}$$
$$x_i = \beta_0 + \beta_1 t_i + \beta_2 ((t_i - t_b)^+)^3 + \epsilon_i \tag{6.1c}$$

Similarly to the trilinear model fit presented in Section 2.3, here x_i is the ith datapoint in the interval of interest, t_i is the time at which measurement x_i is taken, and ϵ_i is the error at datapoint i; I have assumed these are independent and normally distributed with the same variance so the model can be fitted as a linear model, using the maximum likelihood estimators for β_0, β_1, and β_2. As in the original trilinear fit presented in Chapter 2, the value of t_b which maximises the maximum likelihood is then chosen.

The model presented in Equation 6.1a is a bilinear model with breakpoint t_b. The models presented in Equations 6.1b and 6.1c both assume a linear mean until a breakpoint t_b, after which the the mean respectively decreases as a quadratic or cubic function. These could represent a system in which monomer abundance is initially constant, and then begins to decrease as inflammasome formation commences (I have allowed the initial section of the model to be linearly increasing or decreasing to account for effects such as photobleaching or the results of normalising intensities across frames). The bilinear fit is a straight line after the breakpoint, which approximates an exponential decrease at small times after the breakpoint; this would correspond well to the results from the simplified model. The models which are nonlinear after the breakpoints are approximations for Hill functions of the form $H(t) = \frac{c^N}{c^N + t^N}$ for constants c and N, at small times after the breakpoint; this is a common form for a sigmoid function, and would therefore correspond better to the results for the full model with small c_1, with its initial lag phase. Note that it is important to choose n large enough to allow the models to be fitted, but small enough that these approximations still hold. In practice, a value of $n = 50$ was effective, although similar results can be obtained with other values of n.

Note also that I do not assume that the breakpoint t_b coincides with the first breakpoint t_{b_1} identified for each trace in the trilinear fit established in Section 2.3; t_{b_1} is used here only to establish a suitable window for fitting the breaktime t_b.

From this interpretation, β_0 is the initial intensity of the trace, β_1 is the gradient of the first linear section, and β_2 is a scaling parameter controlling the rate of decrease in the second section. An example showing the interval of interest for one representative trace, comparing the three fits with the trace values, is given in Figure 6.1.

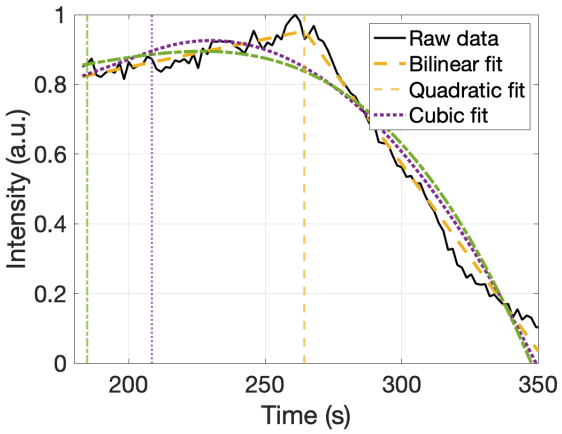


Figure 6.1: Example of results of fitting varying models to the beginning of one representative cytosolic trace in the MitoTracker dataset. The new identified breakpoint t_b for each fit is shown with a vertical line of the same colour. The bilinear model gives the best relative likelihood value.

In order to compare the three models, I fitted each trace and calculated the maximum log-likelihood for that trace. Suppose the maximum log-likelihood for trace i is l_i; then the overall maximum log-likelihood for all traces is the sum of the individual log-likelihoods for each trace, i.e. $\sum_i l_i$. The maximum likelihoods are 5.6042×10^3 for the bilinear model, 5.2111×10^3 for the quadratic model and 5.1795×10^3 for the cubic model. This suggests that the bilinear model is the best fit for the model overall.

In order to directly compare the likelihood of the models, I calculated an AIC value for each model (A_1, A_2 and A_3 for the bilinear, quadratic and cubic respectively); an in-depth explanation of the AIC is given in Section 5.3.3. Then, as shown in [Wagenmakers and Farrell, 2004], we can consider the following statistic for each model:

$$B_i = \exp\left(\frac{min_j(A_j) - A_i}{2}\right) \tag{6.2}$$

We can use B_i to establish a relative likelihood for each model:

$$p_i = \frac{B_i}{B_1 + B_2 + B_3} \tag{6.3}$$

We can use this quantity to determine whether any of the AIC values are significantly different to the others [Wagenmakers and Farrell, 2004]. If any one p_i is larger than the others, this indicates that the corresponding model is most likely to be correct. In this case, we have $p_1 \approx 1$, $p_2 = 1.8945 \times 10^{-171}$ and $p_3 = 3.6573 \times 10^{-185}$. This suggests that the bilinear model is extremely likely to give the best fit of the three models.

Overall, this analysis suggests that there is no lag phase at the onset of ASC depletion, suggesting that the inflammasome behaves similarly to the simplfied model. Otherwise, if

the full model is used, we have c_1, c_3 and p lying in the region defined in Section 5.2.4: $c_1 > 0.546 + 0.385c_3 + 0.209p + 3.60 \times 10^{-3}c_3^2 - 3.46 \times 10^{-3}c_3p + 9.48 \times 10^{-4}p^2$. Note that this region was defined for the parameters c_1, c_3 and p after they have been scaled by the parameter c_2; in the model without scaling, the defined region is as follows:

$$c_1 > 0.546c_2 + 0.385c_3 + 0.209p + 3.60 \times 10^{-3}\frac{c_3^2}{c_2} - 3.46 \times 10^{-3}\frac{c_3p}{c_2} + 9.48 \times 10^{-4}\frac{p^2}{c_2} \quad (6.4)$$

It is particularly interesting that this conclusion has arisen from data measuring the formation of the NLRP3 inflammasome since, in this case, filament intiation occurs via PYD-PYD interactions, while we would expect branching to occur via CARD-CARD interactions; therefore, the lack of lag phase does not arise from the fact that branching and initiation arise from the same protein-protein interaction.

Furthermore, we can use the breakpoint identified t_b as the time at which ASC monomer depletion begins following infection. We can therefore examine the distribution of time taken for preliminary cellular responses to conclude, and for inflammasome formation to begin in earnest (Figure 6.2). There is no significant difference between the FLICA and MitoTracker datasets, or between cytosolic and nuclear traces (Kruskal-Wallis test with null hypothesis that distribution of cytosolic traces and nuclear traces for both FLICA and MitoTracker datasets have the same median, $p = 0.4852$). This is surprising, since it does not agree with the primary analysis of the experimental data in Chapter 2 (which suggests that onset of ASC monomer depletion occurs later in the nuclear region than in the cytosolic region); however, this is more in line with the results from both mathematical models.

6.3 Comparison of experimental data with simplified stochastic model

In Chapter 4, I derived the probablity distributions of the monomer depletion characteristic times $t_{1/q}$ and the single cluster emergence times T_S in the simplified model (see Figure 6.3 for an overview of the simplified model). The monomer depletion characteristic times are directly comparable to the characteristic times $t_{1/q}$ derived from experimental data, while the single cluster emergence time can be compared to the time of inflammasome emergence derived from experimental data (in practice, the second breakpoint t_{b_2} for the trilinear fit in the speck region was used). Prior to analysing either the values of $t_{1/q}$ or the time of

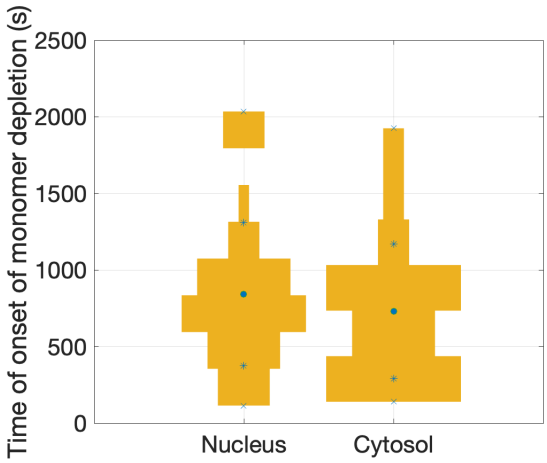
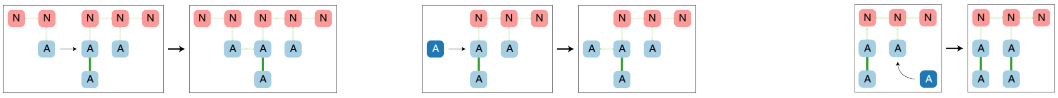


Figure 6.2: Distributions of onset times for monomer depletion in experimental data determined using bilinear breakpoints. The mean value is shown with a blue circle, and the mean $\pm$ standard deviation is shown with a blue cross. There is no significant difference between nuclear and cytosolic compartments; this is in line with the interpretation in the mathematical models.



Reaction C2 (Cluster joining)

Reaction C (Filament branching)

Reaction P (Filament extension)

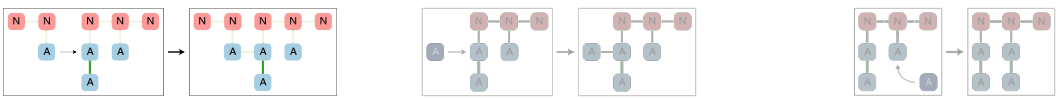
Figure 6.3: The simplified inflammasome model revisited

inflammasome emergence derived from the experimental dataset, in all cases the onset times established in Section 6.2 were subtracted to remove variability arising from the time before monomer depletion begins.

Time of inflammasome formation

Let us consider first the time of single cluster emergence (see Figure 6.4 for an overview of the relevant subsystem of the simplified model).

As demonstrated previously, in the simplified stochastic model, the time of single cluster emergence is hypoexponentially distributed with rates reliant on the parameters c_2, which controls the rate at which clusters join, and X_0, the initial number of NLR oligomers. We can



Reaction C2 is included

Reaction C does not affect this subsystem

Reaction P does not affect this subsystem

Figure 6.4: The simplified inflammasome model with filament extension only revisited

fit the distribution to the data for the time of inflammasome emergence by finding the values of c_2 and X_0 which maximise the likelihood function.

The likelihood of observing a set of times $\{t_i\}$ for $i = 1, 2, \cdots, n$ is as follows:

$$L = \prod_{i=1}^{n} f_{T_s}(t_i) \tag{6.5a}$$

$$f_{T_s}(t_i) = \frac{c_2}{2}(X_0 - 1)!\, X_0! \sum_{j=1}^{X_0-1} e^{-c_2 j(j+1)t_i/2} \frac{(-1)^{j-1} ij(j+1)(2j+1)}{(X_0 - j - 1)!\,(X_0 + j)!} \tag{6.5b}$$

Maximising L is equivalent to maximising the log-likelihood:

$$l = \sum_{i=1}^{n} \log(f_{T_s}(t_i)) \tag{6.6}$$

We take t_i as the set of inflammasome emergence times calculated in Chapter 2 with the depletion onset times from Section 6.2 subtracted. As demonstrated in Section 4.3, The value of X_0 has fairly little effect on the resulting distribution of times of single cluster emergence, so the analysis can be simplified by considering only a limited selection of representative values; I took $X_0 = 10, 50, 100, 500, 1000$.

I used the Matlab function `fminsearch` to find the minimum value of $-l$ for each of these values of X_0 to give the maximum likelihood estimator for c_2.

The largest likelihood occurred when $X_0 = 10, c_2 = 7.56 \times 10^{-3}$ (see Table 6.1 for the full set of values, and Figure 6.5 for a plot of the likelihood functions for different values of X_0 and c_2). This suggests that these values give the best approximation for the values of these parameters, based on the values of X_0 tested. In particular, it is notable that the smallest tested value of X_0 yields the best fit, suggesting that the initial number of NLR oligomers may be very small.

Characteristic times for monomer depletion

A similar analysis can be carried out for the characteristic times $t_{1/q}$. We can use the value of X_0 derived by fitting the single cluster emergence times to determine f_0 by setting $f_0 = 10X_0$. assuming there are ten monomers in each oligomer. For ease, I have focused on the values of $t_{1/2}$ and $t_{1/10}$ for the cytosolic compartment in the experimental data. Once again, I subtracted the fitted onset times derived in Section 6.2 prior to analysis.

The established probability density function for $t_{1/q}$ from Chapter 4 involves taking the matrix exponential of an extremely large matrix (specifically, $N(N+1) \times N(N+1)$) which

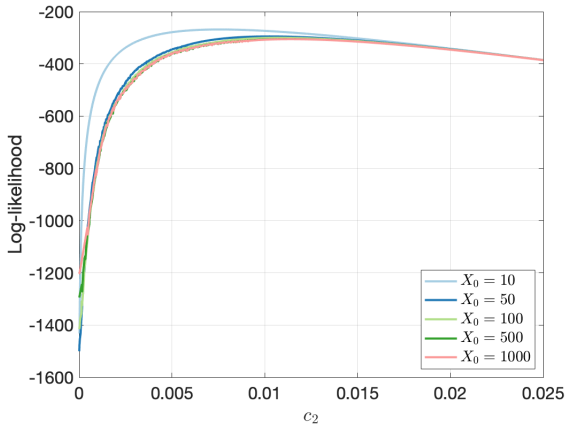
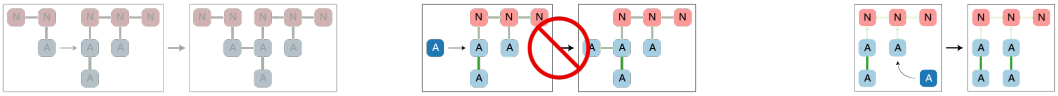


Figure 6.5: Likelihoods for single cluster emergence time derived from analytical results for simplified model while varying the initial number of oligomers X_0 and the cluster joining rate c_2. The largest likelihood occurs for $X_0 = 10$, suggesting a small initial number of NLRs.



Reaction C2 does not affect this subsystem | Reaction C is not included | Reaction P is included

Figure 6.6: The simplified inflammasome model with filament extension only revisited

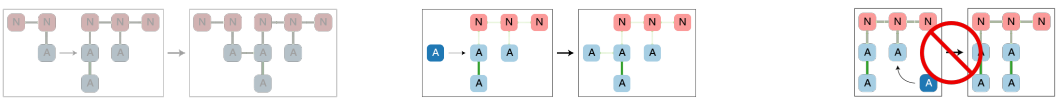
cannot be calculated in an amount of time which is suitable for maximisation. Therefore, I have focused on fitting the two extreme models derived in Section 4.3.2: the model with branching only, and the model with filament extension only (see Figures 6.6 and 6.7 for an overview of the relevant subsystem of the simplified model).

The model with no branching has a log-likelihood function with a relatively simple closed form:

$$l = n\left(\log(p) + \log(f_0) + \log(N) + \log\binom{\hat{x}_0}{N}\right) - pf_0\hat{x}_0\sum_{i=1}^{n} t_i + (N-1)\log\left(\sum_{i=1}^{n}(e^{pf_0t_i} - 1)\right) \tag{6.7}$$

where t_i are the experimentally observed values of $t_{1/q}$.

The form of the likelihood for the model with branching only is somewhat more complex (especially in the case that the rates are not unique), and the full model cannot be fitted



Reaction C2 does not affect this subsystem | Reaction C is included | Reaction P is not included

Figure 6.7: The simplified inflammasome model with branching only revisited

efficiently to within machine precision. However, the form of this function becomes much simpler if we consider only the largest contributing terms to the likelihood function, which correspond to the smallest rates in the hypoexponential distribution. In particular, I calculated the smallest 10 values of $\lambda_i = c(f_0 + i)(\hat{x}_0 - i)$ with i ranging between 0 and $N - 1$, calling these smallest values $\mu_1, \mu_2, \cdots, \mu_{10}$.

The log-likelihood is then as follows:

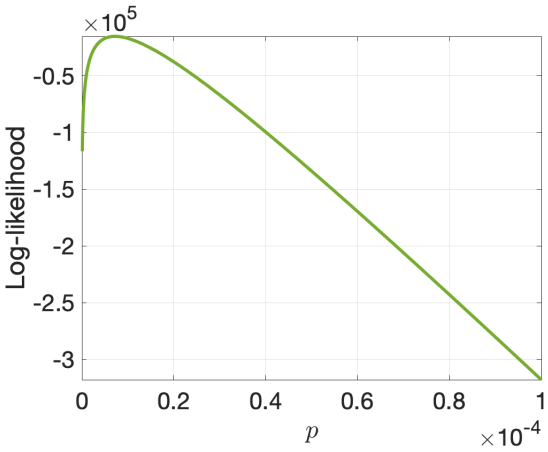
$$l = \sum_i f_{t1/q}(t_i) \tag{6.8a}$$

$$f_{t_{1/q}}(t_i) = \sum_{j=1}^{10} \mu_j e^{-mu_j t_i} \prod_{k=1, k\neq j}^{10} \frac{\mu_k}{\mu_k - \mu_j} \tag{6.8b}$$

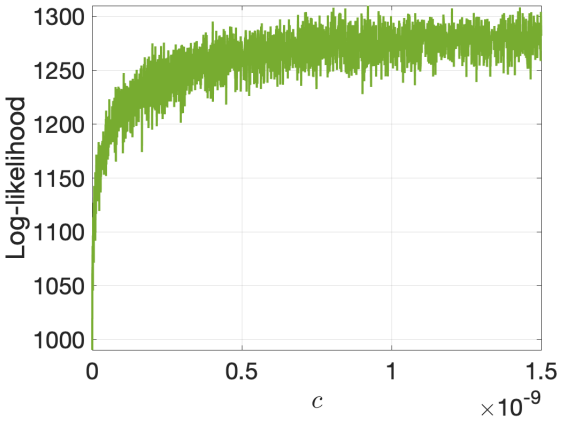
Once again, I minimised the log-likelihood assuming $\hat{x}_0 = 1000$ using the `fminsearch` function for both the model without branching and the model with branching only; the fitted values are given in Table 6.1. Notably, the likelihoods for the model with branching only are considerably higher than the likelihoods for the model with no branching; this suggests that branching may play a key role in inflammasome formation, and that ASC monomer depletion is unlikely to occur via PYD-PYD filament extension reactions only. However, it is worth bearing in mind that the dataset is reasonably small, which may affect the accuracy of the results. Furthermore, the likelihood function oscillates considerably for the model with branching only compared to the model without branching (Figure 6.8); this suggests that the maximisation step may not be reliable, since it may have located a local (rather than global) maximum, although using different initial values for the `fminsearch` function gives reasonably consistent results in terms of order of magnitude. Finally, this result does not suggest that inflammasome formation occurs via branching only, but rather that this is a better fit for the experimental data than the model without branching; this suggests that branching must play a role in inflammasome formation.

6.4 Comparison of experimental data with full stochastic model

We will now turn our attention to the full inflammasome model once more, carrying out a similar model fitting exercise to the process used for the simplified model, paying particular attention to whether fitting the two models gives similar results. A summary of the full model is given in Figure 6.9.

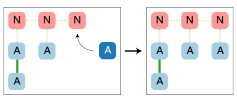


(a) No branching

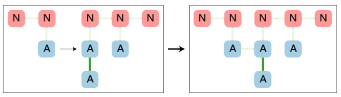


(b) Branching only

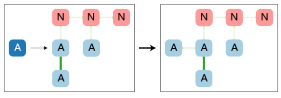
Figure 6.8: Likelihood plots of single cluster emergence time for models with no branching and branching only, from analytical solutions for the simplified model, for an initial number of monomers $\hat{x}_0 = 1000$. The maximum likelihood value is much larger for the model with branching only, although the function oscillates considerably more, which may lead to inaccuracies in fitting.



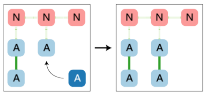
Reaction C1 (Recruitment of ASC monomer by NLR in oligomer)



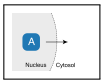
Reaction C2 (Cluster joining)



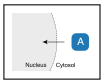
Reaction C3 (Branching of ASC filaments)



Reaction P (ASC filament extension)



Reaction M1 (Movement of ASC monomer from nucleus to cytosol)



Reaction M2 (Movement of ASC monomer from nucleus to cytosol)

Figure 6.9: The full inflammasome model revisited

Fitted parameter	Fitted parameter values	Likelihood
T_s	$X_0 = 10, c_2 = 7.56 \times 10^{-3}$	3.35×10^{-117}
	$X_0 = 50 : c_2 = 1.03 \times 10^{-2}$	9.38×10^{-129}
	$X_0 = 100 : c_2 = 1.08 \times 10^{-2}$	5.01×10^{-133}
	$X_0 = 500 : c_2 = 1.13 \times 10^{-2}$	4.96×10^{-133}
	$X_0 = 1000 : c_2 = 1.14 \times 10^{-2}$	2.71×10^{-133}
$T_{1/10}$, no branching	$p = 7.21 \times 10^{-6}$	$\exp(-1.5289 \times 10^4)$
$T_{1/2}$, no branching	$p = 4.35 \times 10^{-5}$	$\exp(-7.30 \times 10^4)$
$T_{1/10}$, branching only	$c = 1.003 \times 10^{-22}$	$\exp(-1.14 \times 10^3)$
$T_{1/2}$, branching only	$c = 1.003 \times 10^{-22}$	$\exp(-1.33 \times 10^3)$

Table 6.1: Fitted parameter values and likelihoods from comparison of analytical results from the simplified model with initial number of monomers $\hat{x}_0 = 1000$ with experimental data. Fitting the time of single cluster emergence suggests that setting the initial number of NLR oligomers $X_0 = 10$ is the best fit for the data. Fitting $T_{1/2}$ and $T_{1/10}$ gives higher likelihoods for the model with branching only, compared to the model without branching, suggesting that the former is a better conceptual model of inflammasome formation.

In order to assess whether the full model was a good fit for the simulated data, it was not possible to use maximum likelihood estimation, since closed forms of probability density functions for characteristic times had not been obtained; instead, it was necessary to compare the distribution of each characteristic time from the experimental dataset with the corresponding distribution for simulated data representing a number of different variants on the full model, for sample parameter sets (see Section 5.3). In particular, I compared both the single cluster emergence time and the dominant cluster emergence time calculated by mean with the second breakpoint in the speck region, adjusted as for the simplified dataset by subtracting the fitted onset time derived in Section 6.2. I also compared the distributions of $t_{1/10}$ and $t_{1/2}$ for the cytosolic and nuclear regions in the experimental and simulated data. It is important to note that in the simulated data, time has been scaled by multiplication by the parameter c_1 (which controls the rate of ASC monomer interactions with NLRs). It is therefore necessary to scale the times derived from simulations by an estimate for c_1 in order for them to be comparable with the experimental datasets.

I initially compared the simulated datapoints for each characteristic time with the relevant corresponding dataset derived from the experimental data using a two-sample Kolmogorov-Smirnov test, with the null hypothesis that the two samples are drawn from the same distri-

bution. However, since c_1 is unknown, it was necessary to test a variety of values of c_1 to find which was most appropriate. In particular, for each comparison of distributions, I determined the value of c_1 which minimised the Kolmogorov-Smirnov test statistic (which corresponds to maximising the p-value of the test) by applying the `fminsearch` function to the Kolmogorov-Smirnov statistic output of the `kstest2` function in Matlab, with c_1 as the parameter over which the function was to be minimised.

I carried out this analysis for all simulated datasets - namely, the basic form of the full model, the full model without branching, the full model without cluster joining, and the full model with random initial conditions.

The first question of interest was whether the single cluster emergence time or the dominant cluster emergence time was a better fit for the time of inflammasome formation. Interestingly, in general the time of dominant cluster emergence had a much closer distribution to the distribution of inflammasome emergence time derived from the experimental data compared to the time of single cluster emergence (see Tables 6.3 and 6.2 for examples of the best fits). This suggests that dominant cluster emergence time may be a better reflection of inflammasome emergence than single cluster emergence time. The best fit arose for the basic model, suggesting that this might be the best model of inflammasome formation of those tested; in particular, these results suggest that ASC filament branching plays a role in inflammasome formation, and that there are multiple NLR oligomers. However, in comparison to the analysis carried out in the previous section, the fitted value of X_0 is much larger than the value of 10 suggested by analysis of the simplified model.

One clear issue is that the fitted values of c_1 obtained by fitting the dominant cluster emergence times are much smaller than the values obtained for the other characteristic times (single cluster emergence time, and $t_{1/10}$ and $t_{1/2}$ for both nuclear and cytosolic regions). In particular, the parameter values fitted for dominant cluster emergence time were all around 10^{-11}, while the values for the other characteristic times were all between 10^{-4} and 10^{-8}. This may suggest that m_1 is comparatively much smaller than any of the sampled parameter values. As demonstrated in Section 5.2, the time of dominant cluster emergence is primarily dependent on c_3 and X_0, and is broadly independent of m_1, while all other characteristic times do show a dependence on m_1. It is therefore possible that m_1 is much larger relative to the other parameters than any of the values tested using simulations; increasing the value of m_1 would have little effect on the dominant cluster time, but all other characteristic times

Model	$c_1 \times 10^6$	$c_2 \times 10^6$	$c_3 \times 10^6$	$p \times 10^6$	$m_1 \times 10^5$	$m_2/m1$	X_0	**p-value**
Basic	1.50	1.74	2.06	4.85	390	1.44	787	0.688
No branching	2.13	1.10	1.06	N/A	2.56	0.928	597	0.516
Random ICs	1.75	2.58	2.68	9.86	124	1.80	153	0.958

Table 6.2: Fitting simulated results for the single cluster emergence times derived from Latin hypercube simulated data (100 simulations for 100 parameter sets); details of parameters giving rise to simulated distributions of single cluster emergence times for the model variant which is closest to the distribution of inflammasome emergence times derived from experimental dataset. The closest match arises from the Random ICs model.

Model	$c_1 \times 10^{11}$	$c_2 \times 10^{11}$	$c_3 \times 10^{11}$	$p \times 10^{11}$	$m_1 \times 10^5$	$m_2/m1$	X_0	**p-value**
Basic	8.15	50.1	14.7	22.0	1.28	4540	407	0.998
No branching	6.95	55.3	29.8	N/A	7.81	0.157	597	0.993
Random ICs	7.35	8.53	10.1	23.8	191	1.44	787	0.975

Table 6.3: Fitting simulated results for the dominant cluster emergence times derived from Latin hypercube simulated data (100 simulations for 100 parameter sets); details of parameters giving rise to simulated distributions of dominant cluster emergence times for the model variant which is closest to the distribution of inflammasome emergence times derived from experimental dataset. The closest match arises from the basic model.

Model	$c_1 \times 10^6$	$c_2 \times 10^6$	$c_3 \times 10^6$	$p \times 10^6$	$m_1 \times 10^5$	$m_2/m1$	X_0	**p-value**
Basic	5.95	0.855	8.71	0.850	384	0.315	380	0.476
No branching	7.93	63.0	34.0	N/A	8.91	0.157	597	0.478
Random ICs	8.79	3.03	67.0	1.86	115	0.619	314	0.480
One oligomer	1.87	N/A	0.200	9.08	4.96	0.834	N/A	0.478

Table 6.4: Fitting simulated results for the single cluster emergence time and monomer depletion characteristic times derived from Latin hypercube simulated data (100 simulations for 100 parameter sets); details of parameters giving rise to simulated multivariate distributions across characteristic times for each model variant which are closest to the distribution of inflammasome emergence times derived from experimental dataset. The closest match arises from the Random ICs model. P-values are given for the Cramer-von Mies test.

would increase.

Since the fitted values of c_1 were so different to those for the dominant cluster emergence time compared to the other characteristic statistics, I focused on the other characteristic times from this point onwards. In order to ascertain which parameter set gave the best overall match for all characteristic times (with the exception of dominant cluster emergence time), I used a multivariate non-parametric two-sample Cramer-von Mies test developed by Baringhaus and Franz [Baringhaus and Franz, 2004], which tests the null hypothesis that two multivariate samples are drawn from the same multivariate distribution; this could be regarded as an extension to the KS test for multivariate data. The test has been implemented as a package `cramer` [Franz, 2019], in the R statistical software environment [R Core Team, 2016].

I applied the `cramer.test` function to T_s and $t_{1/10}$ and $t_{1/2}$ for both the cytosolic and nuclear compartments for each all simulated results for each parameter set and for each model, comparing with the experimentally-derived values; the exception was the simulated results for the model with one NLR oligomer initially present, for which T_s was excluded. The closest matches occur for the largest p-values; the results for the closest matching parameter set for each model variant are given in Table 6.4. There is fairly little difference between the resulting p-values, although as for the single cluster emergence time alone, the Random ICs model gave rise to the best match.

Note that the parameters giving rise to the closest match to the experimental data are not intended to be a precise estimate for the 'real' parameter values, but are instead are the set of parameters obtained by Latin hypercube sampling that give the results closest to the experimental data. However, the fact that the models with branching included are generally more successful than the corresponding models without branching suggests that branching may be an important factor in inflammasome formation. Furthermore, the most successful models suggest that there is more than one initial NLR oligomer present, suggesting that joining of clusters is another process which should be considered in inflammasome formation. Finally, the reasonably good matches between experimental data and simulated results for the model with random ICs suggests that instead of each NLR oligomer containing exactly 10 monomers, there may be much more variation in oligomer sizes.

It is also worth noting that for all KS and multivariate tests which give the closest results to the experimental data, the parameters do not lie in the region for which there is no lag phase, as defined in Section 5.2.4. This is at odds with the conclusions of Section 6.2, which show a lack of initial lag phase in monomer depletion. The values of X_0 which fit the experimental data best are also much higher than the value of X_0 fitted for the simplified model (Section 6.3), while the value of c_2 derived from the simplified model is much higher than those derived here for the full model. Therefore, the simplified and full models give contradictory results for some aspects of the model. This should be examined in further detail (ideally, with a larger and more consistent experimental dataset) before concrete conclusions can be drawn about the role of ASC recruitment by NLRs at the beginning of the inflammasome formation process.

6.5 Discussion

6.5.1 Summary

In this chapter, I have described a number of approaches to comparing the models of inflammasome formation I have developed with the experimental data described in Chapter 2. I first fitted curves to the early timepoints for each trace derived from the experimental data, to establish whether there is a sharp or gradual initial decrease in ASC monomer abundance following infection. This model fitting suggested that an initial sharp decrease was more likely, and allowed the time taken before ASC monomer depletion occurred to be

determined.

I then used maximum likelihood estimation techniques to fit the characteristic times derived from experimental data to distributions derived for the simplified model in Section 4.3. In particular, this gave an estimate for the cluster joining rate, and suggested that there are initially around 10 oligomers present. This analysis also suggested that filament branching occurs during inflammasome formation, and that cluster growth does not occur purely through filament extension as previous conceptual models have suggested.

Finally, I compared the distribution of the characteristic times from the experimental dataset with the corresponding characteristic times obtained from simulations of the full inflammasome model, both for the basic form of the full models, and for a number of variants, namely, the models without branching, with a single initial NLR oligomer, and with random ICs. While some aspects of the analysis presented here are not entirely consistent with the analysis of the simplified model, they do at least suggest that initially multiple NLRs are present, and that ASC filament branching plays a role in inflammasome formation. Furthermore, analysis of the full model simulations suggests that there may be more variation in NLR oligomer size than the existing conceptual model, in which there are a fixed number of monomers in each NLR oligomer. Finally, the dominant cluster emergence time may give a better reflection of the inflamamsome formation time than single cluster emergence time. This may suggest that inflammasome formation time is dependent on the initial number of NLR oligomers, with smaller numbers of oligomers corresponding to longer inflammasome formation times.

6.5.2 Further work

One of the main shortcomings in the work presented here is the fairly small size, and inconsistent nature, of the experimental dataset analysed, especially as many of the traces derived from the dataset could not be fitted successfully. However, the analysis presented in this chapter could easily be repeated for another similar dataset, which may shed more light on the inconsistencies between the results from fitting the initial portion of the traces (which suggest that the rate of ASC recruitment by NLRs is comparable to the recruitment of ASC by other ASC monomers) and the results from fitting the simplified and full models (which suggest that the rate of ASC recruitment by NLRs is much slower than recruitment of ASC by other ASC monomers). These issues motivate repeating this analysis with a larger

dataset, ideally with much more consistency in time series length and time between frames than in the existing data, as well as a higher signal-to-noise ratio. Increasing the size of the dataset would increase the power of the statistical tests carried out, while decreasing the noise in the dataset would reduce the possibility of incorrect model fits due to random variation discontinuities, and increasing consistency would avoid variation between traces due to avoidable differences between samples. Furthermore, this analysis has only considered the NLRP3 inflammasome; it would be extremely useful to carry out a similar analysis on a system in which a different variety of inflammasome is formed. This would be particularly interesting in cases in which the abundance of NLRs might be different (for example, in the NLRC4 inflammasome); the effect of changing the initial number of NLR oligomers could therefore be directly examined as a variable in the model.

Aside from motivating the production of more data similar to the dataset presented in Chapter 2, the results from this chapter suggest a broader range of areas which would benefit from further general experimental investigation. In particular, the results support the importance of ASC filament branching, and the initial presence of multiple NLR oligomers of different sizes. These could be investigated further using alternative experimental approaches; for example, superresolution imaging of cells with labelled ASC or NLRs during the process of inflammasome formation.

From the mathematical perspective, one shortcoming of the analysis presented in this chapter is that the goodness-of-fit of results for the simplified model cannot be directly compared to the statistical tests carried out on the simulated data for the full model. Thus it is unclear which of the two models gives a better representation of inflammasome formation, and in particular, which aspects of the full model which have been excluded from the simplified model are in fact necessary for an accurate representation of inflammasome formation. The initial analysis carried out in Section 6.2 suggests that, surprisingly, the simplified model may be sufficient to accurately depict monomer depletion at early times, but it is still unclear to what extent the constant joining kernel used in the simplified model is appropriate for this system, and whether the division of the cell into two regions is important in representing the inflammasome formation process. There are two possible approaches which could be used to overcome this drawback; either distributions of characteristic times could be derived for the full model, in a manner similar to the analysis carried out for the simplified model, or the distributions derived for the simplified model could be abandoned in favour of using simu-

lated results. The former approach may be possible, although more technically demanding than the analysis carried out so far (for example, phase-type distributions similar to those presented in Section 4.3 could be developed, although this may require the calculation of the matrix exponential of many large matrices); the latter would potentially be much simpler but, as with the analysis of the full model carried out here, it would still be reliant on the availability of good-quality data to fit.

Finally, as established when analysing the results from the full model, results for the dominant cluster emergence time suggest that a wider range of parameter values should have been considered when running Latin hypercube parameter sampling. This could easily be remedied by carrying out more simulations over a larger range of parameter values, although rerunning simulations over a much wider range of parameter values would be considerably labour-intensive.

Chapter 7

Discussion of intracellular modelling of the inflammasome

In the previous chapters, I have presented a first approach to modelling inflammasome formation on the intracellular level, coupled with analysis of a typical experimental dataset. This research has had a number of key outcomes, the first being the models themselves; this is especially significant since, to my knowledge, there are no existing mathematical or computational models of inflammasome formation with any more sophistication than very simple aggregation mechanisms [Cheng et al., 2010]. The models lay the groundwork for other similar approaches; in particular, by identifying the most important behaviour we would wish an inflammasome formation model to display, and showing how we can integrate structural information about individual protein-protein interactions to gain understanding of the whole inflammasome complex. I have also identified potential problems which may arise from modelling the inflammasome, including the need to use discrete approaches at late times when there are small numbers of reactants present, to avoid biologically-implausible results including clusters which can grow indefinitely in size. The models themselves are also of some mathematical interest, since they contribute to the field of coagulation modelling; this is a novel system which simultaneously includes multiple classes of reactants, movement between spatial compartments, and Smoluchowski-like joining of clusters.

Another key result from this project has been the development of data processing and analysis methodologies, which may be useful beyond this project alone. In particular, the experimental data presented in Chapter 2 is fairly characteristic of imaging data which may be obtained from time series live-imaging of inflammasome formation, and the pipeline de-

veloped here could easily be adapted for use with other experimental datasets. These techniques may be particularly useful since they do not rely on data derived from very intensive, high-resolution but low-throughput techniques such as cryo-EM or NMR imaging, for which live cell imaging is generally not possible, since confocal microscopy provides sufficient resolution for the analysis carried out in this project. Furthermore, this indirect analytical approach is much easier to scale up to analyse larger datasets.

Finally, the results presented in the preceding chapters give an insight into the key processes occurring within the cell during inflammasome formation at a level of detail which would be difficult to achieve through experimental means. This allows us to draw conclusions about the process of inflammasome formation which are currently poorly understood, building towards a truly cohesive and logically consistent conceptual model of the inflammasome formation process. While this project will ideally be only the first step in a continuing cycle of modelling and experimentation of inflammasome formation, already the initial outcomes of this process challenge received wisdom around inflammasome signalling processes, and suggest answers to some of the unresolved questions in this field.

7.1 Towards a conceptual model of inflammasome formation

The results of the model and data analysis suggest a number of key stages in the process of inflammasome formation. Initially, NLRs form oligomers as a response to DAMP or PAMP detection; the results presented here suggest that these oligomers may vary in size, rather than the common assumption (at least for NLRC4 inflammasomes) that all NLR oligomers contain the same number of monomers; this supports the proposed filamentous structures of NLRC4, NLRP1 and AIM2 oligomerisation [Diebolder et al., 2015, Li et al., 2018, Gong et al., 2021, Hollingsworth et al., 2021, Matyszewski et al., 2021]. Furthermore, also contrary to many conceptual models of inflammasome formation, it is likely that there is more than one initial seeding NLR oligomer.

The NLR oligomers then recruit ASC monomers; I have suggested that this process occurs via a chain reaction of ASC 'activation', analogous to the activation of NLRs, but in the case of ASC, allowing the recruitment of further ASC monomers; this model was suggested by the prion-like behaviour of ASC [Cai et al., 2014]. This mechanism explains why aggregation of ASC does not occur spontaneously in endogenous cells. The results presented in

previous chapters also suggest that ASC recruitment and aggregation occurs via both linear, filamentous growth, and the formation of 'branches' in these linear filaments. This challenges the popular view of ASC aggregation, which asserts that ASC recruitment initially occurs exclusively via PYD-PYD interactions, with the resulting filaments crosslinking at later times via CARD-CARD interactions [Broz and Dixit, 2016, Dick et al., 2016, Franklin et al., 2018]; however, it does coincide with the branching ASC filament structures observed by Kuri *et al.* [Kuri et al., 2017]. I have proposed that PYD-PYD interactions lead to filamentous growth, while CARD-CARD reactions are responsible for branching; however, recent studies have suggested that CARD-CARD reactions can also be responsible for linear, filamentous growth [Li et al., 2010, Nambayan et al., 2019, de Alba, 2019, Hollingsworth et al., 2021]. Regardless, CARD-CARD and PYD-PYD interactions are not strictly attached to either mechanism of growth in the mathematical models; rather, the models suggest that there are two methods of growth occurring simultaneously, both of which may be filamentous, but that these two modes of growth are independent; this allows up to two monomers to attach to any growing filament, thus leading to a branch. The models also demonstrate that branching of ASC filaments would allow quicker recruitment of further ASC monomers, which would lead to faster overall formation of the inflammasome once this process has been initiated, since filament branching may lead to a positive feedback loop. Furthermore, I have demonstrated that it is likely that incipient clusters of NLRs and ASC may join partway through formation, since this is necessary for the formation of a single inflammasome.

The analysis presented in previous chapters suggests that as inflammasome formation begins, there is initially a very fast, exponential drop in free ASC monomer counts. This suggests that an initial lag in ASC aggregation does not occur, and that this is not a limiting factor in inflammasome formation. I have also demonstrated that there is a clear separation of ASC monomers in the nuclear and cytosolic regions of the cell, since the inflammasome generally forms outside the nucleus; this causes a slower depletion of ASC in the nuclear region.

Simulations demonstrate the existence of clear phases in the inflammasome formation process, based on the dominant processes at different times and the variation between runs. There is an initial slow phase of growth as NLRs accumulate ASC. Once joining of clusters first occurs, a second phase of rapid growth begins, with a large amount of variation between cells (or simulated experiments). Eventually, one large cluster emerges, which begins to

absorb all smaller clusters. Growth of this cluster begins to slow as ASC monomers and clusters are depleted; at later times, all clusters have been absorbed into a single large cluster. It is important to note that in real biological systems, the process of inflammasome formation takes place concurrently with numerous other cellular processes; in some cases, if the series of phases outlined above takes a particularly long time, the cell may never display the later phases of inflammasome formation simply because other processes have intervened, and the completion of the inflammasome formation process may never occur. In particular, simulated and analytically-derived distributions of dominant cluster emergence time and single cluster emergence time generally have long 'tails' to late times, suggesting that in some cases, some cells would by chance take so long to form inflammasomes that this process is interrupted by other within-cell systems.

The time taken to reach each of the stages outlined above is dependent on different parameters in the model. For example, the time taken to overcome the initial expansion phase before joining occurs is highly dependent on the initial number of NLR oligomers and the rate of ASC recruitment by NLRs; meanwhile, the time taken for a dominant cluster to emerge is dependent on the rate of ASC branching and the initial number of NLRs; and the time of single cluster emergence is dependent on the rate of ASC branching and the rate of movement of monomers from the nuclear to the cytosolic compartment. Interestingly, comparisons of the models to experimental data suggest that dominant cluster formation may be the best fit for inflammasome formation times; the fact that the large-time 'tail' of the distribution of dominant cluster emergence time is reduced as the initial number of NLRs increases may explain why in some systems (for example, NLRC4 inflammasome formation [Man et al., 2014a]) almost all cells form inflammasomes, whereas in others (for example, NLRP3 inflammasomes - see the data presented in Chapter 2) many cells do not form inflammasomes. In the latter case, there may be fewer NLR oligomers present initially, leading to dominant cluster emergence times with much greater positive skewness, and thus theoretically a greater chance of inflammasome formation occurring at times after other cellular processes intervene.

While the framework outlined above is certainly not yet either a definitive or comprehensive model of all processes involved in inflammasome formation, this project has led to the development of useful alternatives to the predominant conceptual models, and presents hypotheses which may be confirmed or rejected on the basis of future experimental work.

7.2 Further work

A clear first stage in building on this project would be to repeat the analysis on further datasets similar to the experimental results presented in Chapter 2. Firstly, it would be useful to repeat the experiments exactly as described in Chapter 2, in order to generate a larger and more consistent dataset, which could be used to establish whether the results presented here are reproducible. It would also be useful to repeat the experiments with other DAMPs/PAMPs, in order to establish the differences between different inflammasome formation processes. This may require some adjustment to the pipeline developed during this project; for example, direct investigation of responses to *Salmonella* infection would require infected cells with *Salmonella* expressing fluorescently-tagged flagellin. In this case, the inflammasome formation process in each individual cell would not begin when the bacteria are introduced to the sample, but rather when that cell is infected. However, it would be possible to establish the time of infection by identifying the time point when the fluorescently-labelled flagellin first appears in the cell. It would also be extremely useful to rerun the analysis in cells in which the relevant NLR has also been fluorescently labelled; this would allow us to interrogate the proposed results relating to NLR oligomerisation presented in earlier chapters.

However, as well as simply rerunning the analysis established in this project with different experimental data, it is also necessary to test the conceptual models proposed above through independent experimentation. One possible approach would be the use of high-resolution techniques such as cryo-EM; in particular, ongoing unpublished work imaging the inflammasome when it is partially formed appears to show branched ASC filament structures similar to those proposed in this project.

As discussed previously, there are also a number of clear limitations to the mathematical models of inflammasome formation proposed here; these could be examined in further iterations of models. There are four main areas in which the models presented here are particularly limited; these are the underlying assumptions of the behaviour of NLRs and ASC, the focus on very early stages of inflammasome formation, the lack of inclusion of other regulatory processes involved in inflammasome formation, and the limited spatial element of the models.

Firstly, in order to construct the model, a number of simplifications were made regarding

ASC and NLR behaviour. In order to explain the lack of ASC aggregation in the absence of any inflammatory stimulus, I assumed that ASC undergoes a change of conformation upon interaction with an activated NLR, which allows it to recruit further ASC monomers. While this is a fairly parsimonious explanation of this biological process which coincides with the prion-like characterisation of ASC, the precise reasons for the lack of spontaneous ASC aggregation are unknown. Naturally, this could be established by further structural analyses of ASC before and after recruitment by NLRs, but alternative hypotheses could also be investigated in a modelling context. One possible such alternative hypothesis could be that ASC can aggregate prior to PAMP/DAMP detection, but that these interactions are highly transient; however, interactions with an NLR oligomer could have a stabilising effect on these aggregates, leading to a reduced dissociation rate.

Another key assumption was that NLR oligomers are initially fully-formed, and do not continue to grow once ASC recruitment begins. This assumption was made since the precise nature of NLR activation and oligomerisation is not entirely clear, and moreover, the method of oligomerisation is different for different PRRs; therefore, by not including the oligomerisation process in the model, it remains applicable to all forms of canonical inflammasome formation. It would be possible to include an input term in the models accounting for the activation of inactive NLRs, although there is currently no existing quantitative data available describing how NLR activation rates change over time. A mathematical alternative would be to investigate a range of different possible functions over time for input of NLR oligomers (for example, constant, decreasing and increasing functions), which could then oligomerise and recruit ASC monomers. This may be able to shed some further light on the complex process of NLR activation without the need for further experimental work. Finally, in some cases the reactions included in the model may be physically impossible due to the geometry of the clusters. This is one reason why intra-cluster reactions were not included in the models created in this research, since space limitations are very likely to affect crosslinking reactions within a cluster. A rigorous representation of the spatial limitations in inflammasome recruitment would rely on including detailed spatial and structural models of the protein involved; this would be laborious and come at the cost of a model which can be investigated analytically. However, an intermediate possibility would be to create a network model of interactions within the inflammasome. The individual proteins could be considered as nodes within a network, with different classes of edges connecting those nodes once CARD-CARD

or PYD-PYD interactions join those proteins; these could be added stochastically over time in a manner similar to the stochastic simulations conducted in Chapter 5. The network properties could then be analysed; in particular, characteristic times such as the time of single cluster emergence (i.e. the time at which the network becomes connected) could be determined for the network and compared with experimental data in a manner similar to the analysis carried out for the models described in previous chapters. Properties of the network could also be used to avoid the inclusion of interactions which might be physically unlikely, for example, tight loops caused by in ASC filaments joining to themselves.

The second limitation of the model presented here is that it is only applicable to the early stages of inflammasome formation, and it includes neither the crosslinking which must occur within inflammasome complex to create the resultant compact ASC 'speck', nor the downstream recruitment of caspase-1. A network model as described above may be useful in addressing the former problem; a metric based on the compactness of the network could be used to assess whether the resultant speck is compact in shape, similar to the endogenous ASC speck, or more reminiscent of filamentous 'Medusa's head' phenotypes [Sahillioglu et al., 2014]. The latter problem could be addressed by adding pro-caspase 1 and caspase-1 as reactants to the models. As well as giving an insight into the involvement of caspase-1 in the inflammasome, the presence of caspase-1 may affect the depletion of ASC in later stages of inflammasome formation, since caspase-1 recruitment to the inflammasome may block further recruitment of ASC monomers (both occur via CARD-CARD interactions).

Another element which could be added to the model is a more detailed account of the regulatory systems controlling NLR activation, for example, COPs and POPs, or proteins involved in NLR PTMs. This was not included in the current model since the process of NLR oligomerisation was not included; furthermore, very little quantitative data is available describing NLR PTMs and interactions with COPs and POPs. A number of points may need to be clarified (for example, the abundance of COPs and POPs before and after DAMP/PAMP detection) before model development could take place, in order to avoid the proposed models being purely speculative.

Finally, a major drawback to the model as presented is that it does not include a spatial element beyond the division of the cell into nuclear and cytosolic compartments; otherwise, I have assumed that the reactants are well-mixed within these compartments. However, within the space of the cell, the need for clusters and monomers to be in spatial proximity in order

to join together is likely to increase the time required for inflammasome formation processes to occur, especially late in the process when there are small total numbers of reactants remaining. It would also be useful to include other subcellular structures which have been implicated in inflammasome formation, including the cytoskeleton — with microtubules and actin having been implicated in NLRP3 and NLRC4 inflammasome formation respectively [Misawa et al., 2013, Man et al., 2014a, Li et al., 2017] — and mitochondria, which have been associated with NLRP3 inflammasome formation [Zhou et al., 2011, Misawa et al., 2013]. The inclusion of the cytoskeleton in particular may direct reactants towards a focal point in which the inflammasome forms, which may explain how reactants come to be in close proximity.

Overall, this research has opened up considerable potential for further modelling in the field of inflammasome signalling; at the broadest possible level, the results presented will be most useful feeding into our understanding of inflammasome formation within the larger context of inflammatory response, both temporally (i.e., considering the downstream effects of cytokine release and cell death), and spatially (i.e. at tissue, organ and organism levels). The models presented here could ultimately form a sub-component of a much broader model, which encompasses these wider effects. These methods could then be the basis for a much wider investigation into the effects of inflammation, and provide a platform which could be used to test possible medical interventions, either to infection, damage or autoinflammation. This will be examined further in the final discussion, Chapter 12.

Chapter 8

Development of a spatial model of lesion formation

8.1 Introduction

Having explored the dynamics of inflammasome signalling in response to infection at the cellular level, we now turn our attention to the larger-scale effects of inflammasome formation and, in particular, the formation of lesions in the liver following *Salmonella* infection.

In order to better understand the processes which define the spatial organisation of lesions in *Salmonella* infection, I have created a spatial PDE model describing the dynamics of infected and uninfected phagocytes as well as bacteria and chemokines over time. In this chapter, I will describe the processes included in the model, as well as the simplifications and assumptions involved. I also describe a basic set of parameters which can be used with this model, which have been derived from experimental literature.

8.2 Lesions model with single phagocyte population

8.2.1 Variables and reactions

I have focused my attention on four variables: the numbers of infected and uninfected phagocytes, as well as the number of extracellular bacteria and the concentration of chemokines released by phagocytes while infected. I track each of these variables over time, and over a spatial domain representing an area of liver tissue. I assume that all phagocytes as well as chemokines and extracellular bacteria diffuse freely across the spatial domain, but that

Variable	Explanation
S	Concentration of uninfected phagocytes
I	Concentration of infected phagocytes
b	Concentration of extracellular bacteria
c	Concentration of phagocyte-attracting chemokine

Table 8.1: Variables used in the spatial lesion model

all phagocytes undergo chemotactic attraction towards areas with a high concentration of chemokines. When in spatial proximity, bacteria may infect uninfected phagocytes. Infected cells may undergo lysis or apoptosis; in the former case, bacteria are released from the cells. Chemokines are degraded and extracellular bacteria die at a constant rate. Although I assume that bacteria reproduce within the cellular environment, I do not explicitly track the number of bacteria in the cell, but simply assume that a constant number N of bacteria are released when a cell is lysed. Finally, I assume that uninfected phagocytes are introduced to the system at a constant rate throughout the domain, and are removed at a linear rate.

The variables and parameters used in the model are given in Tables 8.1 and 8.2 respectively. The reactions in the model are illustrated in Figure 8.1. The resulting PDEs are given in System 8.1.

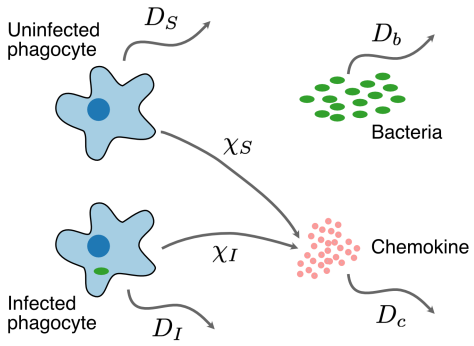
$$\frac{\partial S}{\partial t} = \underbrace{D_S \nabla^2 S}_{\text{Diffusion}} \underbrace{-\nabla \cdot (\chi_S S \nabla c)}_{\text{chemotaxis}} \overbrace{-\beta S \frac{b}{b + b_{1/2}}}^{\text{Infection}} \underbrace{+r_S}_{\text{Input of fresh cells}} \overbrace{-\gamma_S S}^{\text{Removal of cells}} \tag{8.1a}$$

$$\frac{\partial I}{\partial t} = \underbrace{D_I \nabla^2 I}_{\text{Diffusion}} \underbrace{-\nabla \cdot (\chi_I I \nabla c)}_{\text{Chemotaxis}} \overbrace{+\beta S \frac{b}{b + b_{1/2}}}^{\text{Infection}} \underbrace{-(l + a) I}_{\text{Cell death}} \tag{8.1b}$$

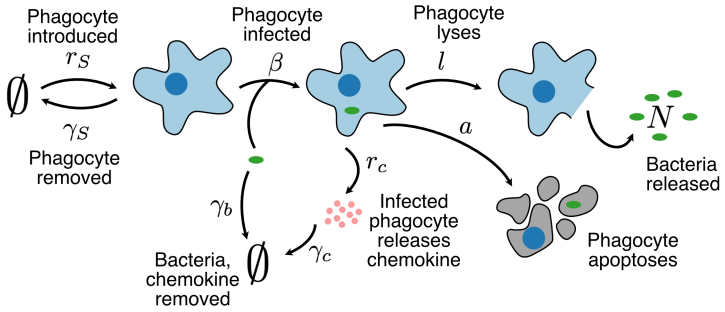
$$\frac{\partial b}{\partial t} = \underbrace{D_b \nabla^2 b}_{\text{Diffusion}} \underbrace{-\beta S \frac{b}{b + b_{1/2}}}_{\text{Infection}} \overbrace{+N l I}^{\text{Bacteria release during lysis}} \underbrace{-\gamma_b b}_{\text{Death of extracellular bacteria}} \tag{8.1c}$$

$$\frac{\partial c}{\partial t} = \underbrace{D_c \nabla^2 c}_{\text{Diffusion}} \overbrace{+r_c I}^{\text{Chemokine production by infected cells}} \underbrace{-\gamma_c c}_{\text{Chemokine decay}} \tag{8.1d}$$

The domain of the system is a two-dimensional area $D = [0, L]^2$, with reflecting boundary conditions; this is a somewhat arbitrary choice which is particularly amenable to analysis, and thus a good choice for a first modelling attempt; considering only a two-dimensional domain further simplifies analysis, although these choices of domain architecture may affect



(a) Movement of reactants in the spatial lesions model. Uninfected and infected phagocytes, bacteria and chemokines diffuse with diffusion coefficients D_S, D_I, D_b and D_c respectively. Uninfected and infected phagocytes undergo chemotaxis up a gradient of chemokine concentration with chemotaxis parameters χ_S and χ_p respectively.



(b) Non-spatial reactions in the spatial lesions model. Uninfected phagocytes are introduced at a constant rate r_S, and are removed from the system at a rate γ_S. Bacteria infect uninfected phagocytes at a rate β. Infected phagocytes produce chemokines at a rate r_c. Infected cells lyse and apoptose at rates l and a respectively. Cells which lyse release N bacteria. Bacteria and chemokines are removed at rates γ_b and γ_c respectively.

Figure 8.1: Reactions considered in the spatial model of lesion formation.

Parameter	Explanation
D_S, D_I, D_b, D_c	Diffusion coefficients
χ_S, χ_I	Chemotaxis rates
β	Bacterial infection rate
$b_{1/2}$	Half-saturation concentration for bacterial infection
r_S	Rate of introduction of uninfected phagocytes
γ_S	Rate of removal of uninfected phagocytes
l	Lysis rate of infected phagocytes
a	Apoptosis rate of infected phagocytes
N	Average bacterial release following lysis
γ_b	Bacterial death/removal rate
r_c	Chemokine production rate
γ_c	Chemokine decay/removal rate

Table 8.2: Parameters used in the spatial lesion model

the structure of the solutions.

The intial conditions are $S(\boldsymbol{x}, 0) = S_0(\boldsymbol{x})$, $p(\boldsymbol{x}, 0) = 0$, $b(\boldsymbol{x}, 0) = b_0(\boldsymbol{x})$ and $c(\boldsymbol{x}, 0) = 0$ for some functions $S_0(\boldsymbol{x})$ and $b_0(\boldsymbol{x})$. These functions represent the distribution of resident phagocytes and bacteria introduced at $t = 0$, when the infection begins. In general, I have assumed that resident phagocytes are initially evenly distributed across the domain, while bacteria are introduced in a localised area.

8.2.2 Simplifications and assumptions

The model presented here is clearly a very much simplified representation of the biological process of lesion formation. However, I have attempted to identify the key processes driving the dynamics and spatial distribution of the reactants in order to create the most parsimonious possible model, to test whether these processes alone are sufficient for lesion formation; if this is unsuccessful in modelling lesion formation, the model may be adjusted to change the characterisation of the biological processes included, or to include other processes.

Firstly, I have assumed that the domain is spatially homogeneous, and that cells, bacteria and chemokines undergo diffusion freely throughout the domain. I have not considered

the lobule structure of the liver tissue, or interactions with other cells and structures such as hepatocytes and the ECM. This reduces the complexity of the model, and furthermore it is unclear how interactions with other structures and cells affects lesion formation, if at all. Likewise, I have not explicitly modelled the blood supply in the domain, but for simplicity I assume that phagocytes are introduced at a constant rate throughout the domain; this is naturally a considerable simplification of the system, and other similar models include an input term proportional to the local concentration of chemokines [Dunster et al., 2014, Bayani et al., 2020a]. However, the constant input of phagocytes has two advantages: firstly, in the absence of bacteria, infected phagocytes, and chemokines, the system remains at an equilibrium with $S = r_S/\gamma_S$; unlike a model in which phagocyte input in a given area is proportional to local chemokine concentration, this would allow a nonzero population of resident phagocytes to persist in the absence of infection. Secondly, if local phagocyte input is proportional to local chemokine concentration, this may limit the input of uninfected phagocytes to areas already containing infected cells; essentially, this may result in a situation in which uninfected phagocytes can only be introduced to the system very close to an existing lesion.

I also assume that the cells, bacteria and chemokines cannot leave the domain, but also that no chemokines or bacteria can enter the domain. Of course, during the course of an infection we would expect bacteria to enter the liver over an extended time period; however, in the interest of simplicity and parsimony I have not considered this possibility here.

Another key simplification in this model is the grouping of phagocytic cells only by their infection status; namely, I do not distinguish between neutrophils and macrophages. This is justified since the initial neutrophil response is extremely transient, and differentiating between the roles played by the different cells may not be necessary to explain the spatial dynamics of this system. Furthermore, I have not included macrophage activation in this system, since the classification of macrophages based on activation status would require the inclusion of further variables in the model, which would be unlikely to materially affect results beyond a delay as macrophages pass from the inactive to activated class. I have also assumed that cells do not replicate within the timescale of this model, and that infected cells do not die by any other means than programmed cell death following infection.

Likewise, I have amalgamated all chemokines released by infected cells into a single variable; naturally, in the biological system, multiple chemokines are released, and play subtly different roles in driving inflammation including recruitment of other innate immune cells. The

use of a single variable is a parsimonious construction which allows us to focus on the potential role played by chemokines in recruitment of phagocytes, since the focus for this model is understanding the interplay between bacteria and cells, rather than the roles of individual chemokines. Furthermore, I have not included chemokine production from uninfected cells which have come into contact with extracellular bacteria, or release of chemokines following lysis. The total amount of chemokines released by either of these processes is likely to be much less than the chemokines released by infected cells during their lifetimes, and furthermore, no quantitative data is available to parametrise these processes; therefore, the omission of these processes is justified. Finally, I have assumed that chemokine binding to cell receptors is transient, and that chemokines are thus not removed from the system by binding; this is justified by the extremely short timescale of these reactions compared to the process of lesion formation.

I have also made a number of assumptions about the behaviour of bacteria. I have assumed that bacteria infect uninfected cells only; this is justified since bacteria are much less likely to infect cells which are already infected [Gog et al., 2012]. I have assumed that infection occurs at a rate proportional to the local concentration of uninfected cells and a Michaelis-Menten type function of the local concentration of bacteria; this reflects saturation in the infection process, and is important in representing limitations in the number of bacteria which can be taken up by cells due to spatial constraints. This form has been used in other models of bacterial infection [Lauffenburger and Kennedy, 1983, Schokker et al., 2013]. Furthermore, I have assumed that bacteria do not reproduce extracellularly, but do reproduce intracellularly, reflecting the preferential conditions for bacteria in the intracellular environment. It would certainly be possible to add a simple linear term to include extracellular bacterial reproduction, but for the sake of model parsimony this was excluded from the current model. Although naturally there is a distribution of numbers of bacteria released upon cell lysis, I assume that a fixed number $N > 1$ of bacteria are released whenever a cell lyses. While it would be possible to introduce a distribution of possible numbers of bacteria released upon lysis rather than a single fixed value, or even to explicitly model intracellular bacterial reproduction, since the number of bacteria released upon lysis is generally extremely low the use of a constant number of bacteria released is justified [Brown et al., 2006, Gog et al., 2012, Man et al., 2014a].

Finally, the use of linear death rates, chemotaxis rates and rate of introductions of new

phagocytes are certainly not a perfect representation of the biological system, but since these rates have not been quantified, this is a logical choice for a first model of this system.

This model contrasts to previous spatial models of inflammatory processes in a number of respects. The innovative aspects of this model can be generally considered either as processes which have been included which are unusual in other models (for example, explicit quantification of bacteria and the inclusion of lysis), or processes which have been simplified since they are of less interest to this research (for example, consolidating different varieties of cells and chemokines into single variables).

Firstly, the focus on bacterial dynamics in this model is fairly unusual, especially in spatial PDE models, which have often been more general in nature rather than specifically focusing on bacterial infections; in particular, it is extremely unusual for models to include lytic cell death, and the release of bacteria following lysis. However, since the balance of lysis and apoptosis are particularly important in understanding the role of the inflammasome, I have included both in this model.

Secondly, in this model, parameters describing movement of phagocytes may be different for infected and uninfected cells. This is an unusual distinction, since most models assume that rates of diffusion and chemotaxis are the same for all cells of the same class, regardless of infection status. However, in examining the downstream effects of inflammasome formation, one area of particular interest is the observed effect of decrease in cellular motility; allowing movement rates to vary between infected and uninfected cells will allow the effect on lesion formation to be examined more systematically.

Thirdly, neutrophils and macrophages are regarded as a single class of cells. While they naturally play different roles in the resolution of infection, for the reasons outlined above, these have been consolidated into a single class of cells for the sake of parsimony.

8.2.3 Expected behaviour of the model

In order to accurately represent the process of lesion formation, we expect the results of our model to show a number of key characteristics. Firstly, it is vital that we see the formation of distinct clusters of infected phagocytes which represent lesions. We would expect these to be static, and roughly spherical in shape (circular in the 2D model). If such solutions were not possible with the model as it stands, this would demonstrate the need for inclusion of another process to stabilise clusters as they form, such as the late arrival of fibroblasts to ‘close off’

clusters. This has been proposed as a hypothetical mechanism for lesion formation.

Secondly, we expect to see lesions which do not grow significantly over time once they have reached a certain size, and that instead the total number of lesions increases as the bacterial population increases.

Finally, we would not expect to see infected macrophages or extracellular bacteria migrating between clusters, but rather infected macrophages remain associated with one lesion until cell death occurs, and that any intracellular bacteria will either remain associated with the lesion they originated from (for example, infecting new macrophages which have been recruited to the lesion) or form new lesions elsewhere.

8.3 Parametrisation

In order to analyse the behaviour of the models described above, it was necessary to fix the values of some parameters in the model to reduce the overall size of the parameter space. In many cases, direct measurements have not been taken experimentally, but suitable estimates can be obtained by using experimentally-derived values for parameters describing similar systems. In particular, I have used parameters calculated for TNF-α as representative chemokine parameters in this model, as well as parameters measured for macrophages for phagocyte parameters. An overview of the parameters is given in Table 8.3.

For their estimate of macrophage diffusion coefficients in their model of atheroscelrosis, Hao and Friedman use an experimentally-derived estimate presented by Kim *et al* for the diffusion coefficient of glioma cells ($10^{-11}cm^2s^{-1}$) [Kim et al., 2011, Hao and Friedman, 2014]. They also use an estimate of chemotactic sensitivity derived by Kim *et al.* ($(1.8-4.2)\times 10^{-7}cm^2g^{-1}s^{-1}$), also originally for glioma cells. These are used as estimates for D_S and χ_S respectively.

Yeh *et al.* calculate values of the diffusion coefficient of macrophages in different cell densities [Yeh et al., 2017]; all results lie in the range $(0.47-0.60)\times 10^{-3}mm^2s^{-1}$, suggesting that the estimate used by Hao and Friedman is reasonably reliable.

For their model of sarcoidosis, Hao *et al.* estimate the rate of production of TNF-α by macrophages to be $2.86\times 10^{-3}d^{-1}$ [Hao et al., 2014]. I have used this as an estimate for r_c.

Man *et al.* demonstrate that following infection, macrophages undergo stiffening and thereafter experience a reduction in motility [Man et al., 2014a]. They suggest that this

may be a strategy to reduce infection. Since the chemotaxis and diffusion rates for infected macrophages and neutrophils have not been formally quantified, I will use this model to investigate the results upon varying these rates in the following chapters. As a baseline, I set D_I and χ_I to be equal to zero. This is not necessarily intended to be an accurate reflection of the processes occurring *in vivo*, but instead gives a particularly extreme solution which can be contrasted with solutions when different parametrisations are used (see Figure 8.2).

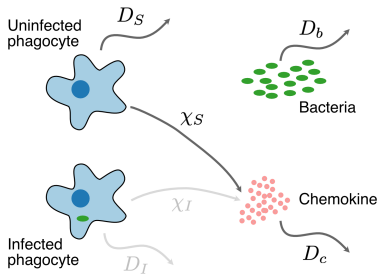
In their examination of *Salmonella* pattern formation on agar, Woodward *et al.* estimate a diffusion coefficient of $2.8 \times 10^{-6} cm^2 s^{-1}$, while Furter *et al.* give an estimate of $0.233 \mu m^2 s^{-1}$ for the diffusion coefficient of bacteria in the colon mucus surface [Woodward et al., 1995, Furter et al., 2019] . Since these are approximately in the same order of magnitude, I use the result of Furter *et al.* as an estimate for D_c.

Ross and Pompano derive estimates for free diffusion coefficients of various cytokines (including human and murine TNF-α), as well as diffusion coefficients for these cytokines in lymph node tissue [Ross and Pompano, 2018]. The measured diffusion coefficients for murine TNF-α in T-cell and B-cell zones are $4.4 \times 10^{-7} cm^2 s^{-1}$ and $5.4 \times 10^{-7} cm^2 s^{-1}$ respectively. Since these are of the same order of magnitude, I use the former estimate for D_c.

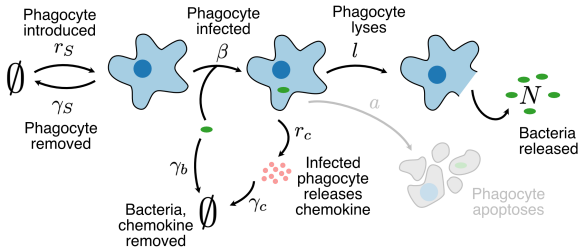
Brown *et al.* establish a constant lysis rate of cells independent of bacterial count, and find the mean number of bacteria released during cell lysis to be 7.16 [Brown et al., 2006]; I use this as an estimate for N.

Schokker *et al.* present a model of *S. enterica* proliferation in chicken intestinal tissue; they use estimates of the infection rate of macrophages by *Salmonella* as $0.1 d^{-1}$, adjusting this term by including a Michaelis-Menten-like factor to reflect saturation of bacteria uptake by cells with half-saturation value $6 \times 10^5 cm^{-3}$ [Schokker et al., 2013]. I use these as estimate for β and $b_{1/2}$ respectively. Schockker *et al.* also use a lysis rate of infected cells of $0.8 d^{-1}$, which is used as an estimate for l, and give a baseline death rate of extracellular *Salmonella* as $27.8 d^{-1}$, which is used as an estimate for γ_b. They also assume that the rate of macrophage influx is approximately 1% of the initial number of macrophages per day, giving an estimate of $3 \times 10^5 cm^{-3} d^{-1}$. I use this as an initial estimate for r_S, although this is further explored in the following chapters.

The role of apoptosis in infected cells is debated; similarly to the movement parameters



(a) Infected phagocyte motion removed



(b) Apoptosis removed

Figure 8.2: Baseline model of lesion formation. Infected phagocyte movement and apoptosis are removed.

for infected cells, I also initially set the apoptosis rate a for infected cells to be zero, not because this is expected to be an accurate reflection of the processes occurring *in vivo*, but rather to give an extreme baseline against which results for other parametrisations can be compared. The results of varying the apoptosis rate are explored in later chapters. A summary of the baseline model of lesions modelling is given in Figure 8.2.

The standard set of parameters used for the model is given in Table 8.3.

I have also used estimates derived from literature to set the initial conditions of the models (the precise spatial distributions will be explained in more detail in Section 10.2).

The density of hepatocytes in pig livers is around $1.47 \times 10^6\ mm^{-3}$ [Junatas et al., 2017]. Hepatocytes account for 60% of the total cell population, while Kupffer cells account for around 10% [Williams and Iatropoulos, 2002]. Thus I used a mean initial concentration of $1.47 \times 10^6 \div 6 \approx 2.45 \times 10^8 cm^{-3}$ for the resident uninfected phagocytes. In the model developed by Schokker *et al.*, the initial density of *Salmonella* in tissues is estimated to be $200 cm^{-3}$; this estimate is also used here.

8.4 Discussion

8.4.1 Summary

In this chapter, I have presented a simple PDE model of lesion formation in the liver during systemic *Salmonella* infection, similar to other models of inflammation in a tissue context. However, unlike other existing models, I have included bacterial dynamics and cell lysis, as well as allowing movement parameters to differ for infected and uninfected cells, in order to examine the downstream effects of inflammasome formation. I have described in-depth the

Parameter	Value used	Unit	Source
D_S	8.64×10^{-7}	cm^2d^{-1}	[Hao and Friedman, 2014], [Hao et al., 2014]
D_I	0	cm^2d^{-1}	[Man et al., 2014a]
D_b	2.01×10^{-5}	cm^2d^{-1}	[Furter et al., 2019]
D_c	3×10^{-2}	cm^2d^{-1}	[Hao et al., 2014],[Ross and Pompano, 2018]
χ_S	1.61×10^{-2}	$cm^4g^{-1}d^{-1}$	[Hao and Friedman, 2014], [Kim et al., 2009]
χ_I	0	$cm^5g^{-1}d^{-1}$	[Man et al., 2014a]
β	0.1	d^{-1}	[Schokker et al., 2013]
$b_{1/2}$	6×10^5	cm^{-2}	[Schokker et al., 2013]
r_S	$3times10^5$	$cm^{-2}d^{-1}$	[Schokker et al., 2013]
γ_S	0.03	d^{-1}	[Dunster et al., 2014]
γ_I	0	d^{-1}	[Man et al., 2014a]
l	0.8	d^{-1}	[Schokker et al., 2013]
a	0	d^{-1}	NA
N	7.16	NA	[Brown et al., 2006]
γ_b	27.8	d^{-1}	[Schokker et al., 2013]
r_c	2.86×10^{-12}	d^{-1}	[Hao et al., 2014]
γ_c	55.45	d^{-1}	[Hao et al., 2014]

Table 8.3: Parameter values used in the spatial PDE model of lesion formation

underlying assumptions and simplifications in this model, as well as criteria for the solutions of such a model to be regarded as successful. Finally, I have also established a 'standard' set of parameters which can be used as a baseline for this model, based on experimental estimates. In some cases, parameters are not explicitly known, but the results from varying these parameters will be investigated in future chapters.

8.4.2 Further work

Naturally many simplifications and assumptions have been made to create a model which balances parsimony and tractability with inclusion of the main important processes in this system. As a first model, I have aimed for the simplest possible summary of the processes involved; if this model is shown to be non-optimal, this initial model can then be revised in the ongoing process of model refinement. Nonetheless, there are some obvious areas which could be altered in future iterations of this model.

In particular, a clear next step would be to introduce spatial heterogeneity to the domain, for example, incorporating the tissue structure of the liver by creating separate spatial compartments representing lobules of the liver, or defining the location of blood vessels within the domain. This would require only alteration of the domain on which the model is defined, and the boundary conditions. However, for the initial modelling process, I decided to use a simple square domain in order to establish whether interactions with a more complex boundary were necessary for lesion formation.

It would also be relatively simple to include more different classes of cells to the model; for instance, differentiating between neutrophils and activated and inactive macrophages. This could be achieved similarly to the models presented by Dunster *et al.* and Bayani *et al.* [Dunster et al., 2014, Bayani et al., 2020a]. This would be a reasonably easy change to introduce, but would mostly be important if we were particularly interested in the difference in behaviour between the subdivisions of the existing variables, which is not the case for this research.

There are also a number of processes which are not currently included in the model, which could be added. For example, we could include a source term similar for that of input of new systemic phagocytes representing the arrival of new bacteria over time. Likewise, we could use a more biologically accurate function for this source term rather than the constant input term used in the model presented here. However, we currently know reasonably little

about the temporal aspects of this process; it might be particularly interesting to consider this model embedded within a larger model tracking the movement of bacteria and macrophages between organs in the body. This need only be fairly simple, with a few key organs represented as separate compartments as in the model developed by Grant *et al.* modelling the movement of different bacterial populations between the blood, liver and spleen, combined with a branching process model of bacterial population growth [Grant et al., 2008a]. A similar simple within-host model could potentially be coupled with the spatial model presented here to give more of a organism-level view of the effects of bacterial replication within lesions.

An alternative to the assumption that upon lysis all infected cells release exactly N bacteria would be to explicitly include bacterial reproduction within the cell. The class of infected cells could be divided into subclasses $I_1, I_2, I_3, \cdots$ where I_n is the concentration of cells containing exactly n bacteria. This could potentially also be coupled with allowing bacteria to reproduce extracellularly, for example, using logistic growth. While this would be an extremely interesting extension to the model, the large number of variables introduced greatly increases the complexity of the model, so in this project I have confined my interest to at most two classes of infected cells.

Finally, we could allow bacteria to infect cells which have already been infected; analysis by Gog *et al.* has already established the reinfection rates of cells which have already been infected in *S. Tymphimurium* infection [Gog et al., 2012], so this would certainly be feasible, although I did not consider this completely necessary for a first model. Likewise, the dynamics of chemokine release could be made more representative of the biological system by the inclusion of the release of chemokines during lysis and by uninfected cells which have been in close contact with bacteria (for example, in a manner analogous to macrophage activation in the model of Dunster *et al.* [Dunster et al., 2014]). However, it would be useful to have more experimental data in order to quantify the respective rates of chemokine release through these processes.

It is worth bearing in mind that some of the limitations in this modelling process are due to the lack of availability of formal quantitative data describing this system. Many of the parameter estimates have been taken from other, similar systems; much of the experimental data on lesion formation consists of two-dimensional images of histological samples, which in general have not been subjected to a particularly sophisticated level of quantitative analysis. This therefore demonstrates that this field has great future potential for further collaboration

between modellers and experimental biologists.

Chapter 9

Analysis of the spatial model of lesion formation

9.1 Introduction

While the models presented in Chapter 8 are relatively simple examples of spatial PDE models, it is still not possible to find analytic solutions. Although I will present numerical solutions to the models in Chapter 10, it is nonetheless possible to gain some insight into the behaviour of these models using analytical approaches. In particular, I will examine parametrisations of the model for which spatial patterning can occur; these naturally-occurring spatial patterns may be the reason for stable lesion formation. I will use methods similar to those established by Turing [Turing, 1952], and used in other models of inflammatory response (for example, [Keller and Segel, 1970, Penner et al., 2012, Bayani et al., 2020a]).

In particular, I will investigate how varying parameters corresponding to processes of particular interest (apoptosis, and the motility of infected cells) affects the stability of the solutions, and establish whether spatial patterning can occur. I will also demonstrate that the stability of the solution is sensitive to the rate of input of uninfected phagocytes into the system (a parameter for which we do not have a well-established value estimate). I will demonstrate that spatial patterning is possible only for particular combinations of values for these parameters, and that the formation of spatial patterning is particularly dependent on the rates of apoptosis and uninfected cell influx. I will also show that increasing apoptosis rate increases the overall stability of the model, while increasing the influx of uninfected cells or the motility of infected cells has a destabilising effect.

Finally, this analysis allows us to establish estimates for the approximate size of structures formed during spatial patterning as a function of the parameters in the model; this could be compared to experimental data to establish which parameter sets are most likely to lead to lesion-like spatial patterns. Throughout this chapter, I will use the baseline parameter set established in Table 8.3.

9.2 Stability analysis

9.2.1 Simplified model

We will consider the spatial PDE model of lesion formation presented in the previous chapter:

$$\begin{aligned}
\frac{\partial S}{\partial t} &= D_S \nabla^2 S - \nabla \cdot (\chi_S S \nabla c) - \beta S \frac{b}{b + b_{1/2}} + r_S - \gamma_S S \\
\frac{\partial I}{\partial t} &= D_I \nabla^2 I - \nabla \cdot (\chi_I I \nabla c) + \beta S \frac{b}{b + b_{1/2}} - (l + a) I \\
\frac{\partial b}{\partial t} &= D_b \nabla^2 b - \beta S \frac{b}{b + b_{1/2}} + N l I - \gamma_b b \\
\frac{\partial c}{\partial t} &= D_c \nabla^2 c + r_c I - \gamma_c c
\end{aligned} \tag{8.1 revisited}$$

In order to establish when spatial patterning is possible, we are looking for a possible state which is spatially inhomogeneous but stable over time. This may arise when a spatially homogeneous system is perturbed away from a stable steady state, and transport terms (diffusion and chemotaxis) destabilise the system. We can demonstrate that such an instability can occur in solutions to our lesion by stability analysis; we first identify a stable steady state in our system in the absence of transport terms, and then demonstrate that once transport terms are re-introduced some perturbations may grow with time, allowing patterns to emerge.

Note that, throughout this chapter, I will assume that valid parametrisations of the model will have all parameters taking non-negative values. Solutions are valid when S, I, b and c are non-negative for all time.

We write solutions in the form $\boldsymbol{u}(\boldsymbol{x}, t) = (S(\boldsymbol{x}, t), I(\boldsymbol{x}, t), b(\boldsymbol{x}, t), c(\boldsymbol{x}, t))$.

Removing the transport terms gives the following system:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta S \frac{b}{b + b_{1/2}} + r_S - \gamma_S S \tag{9.1a}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta S \frac{b}{b + b_{1/2}} - (l + a) I \tag{9.1b}$$

$$\frac{\mathrm{d}b}{\mathrm{d}t} = -\beta S\frac{b}{b+b_{1/2}} + NlI - \gamma_b b \tag{9.1c}$$

$$\frac{\mathrm{d}c}{\mathrm{d}t} = r_c I - \gamma_c c \tag{9.1d}$$

The spatially homogeneous system has a unique nontrivial fixed point at $\boldsymbol{u}^* = (S^*, I^*, b^*, c^*)$ where S^*, I^*, b^*, c^* are as follows:

$$S^* = \frac{r_S(Nl-(l+a)) + b_{1/2}\gamma_b(l+a)}{(\beta+\gamma_S)(Nl-(l+a))} \tag{9.2a}$$

$$I^* = \frac{\beta r_S(Nl-(l+a)) - b_{1/2}\gamma_b\gamma_S(l+a)}{(\beta+\gamma_S)(a+l)(Nl-(l+a))} \tag{9.2b}$$

$$b^* = \frac{\beta r_S(Nl-(l+a)) - b_{1/2}\gamma_b\gamma_S(l+a)}{\gamma_b(\beta+\gamma_S)(l+a)} \tag{9.2c}$$

$$c^* = \frac{r_c(\beta r_S(Nl-(l+a)) - b_{1/2}\gamma_b\gamma_S(l+a))}{\gamma_c(\beta+\gamma_S)(a+l)(Nl-(l+a))} \tag{9.2d}$$

This fixed point corresponds to a steady state in the homogeneous system, in which bacteria and infected phagocytes exist in equilibrium with macrophages and chemokines. Note that there is also a disease-free steady state $S = r_S/\gamma_S, I = 0, b = 0, c = 0$; this is the equilibrium in which the system exists in the absence of bacteria or infected phagocytes. The non-zero value of S here may correspond to the resident macrophage population.

The nontrivial fixed point $\boldsymbol{u}^*$ exists when the following conditions hold:

$$Nl > l + a \tag{9.3}$$

$$b_{1/2}\gamma_b\gamma_S(l+a) < \beta r_S(Nl-(l+a)) \tag{9.4}$$

We can apply these conditions to varying a and r_S, the two parameters we wish to investigate further. In particular, if we take all other parameters as in the standard parametrisation (Table 8.3), the conditions become:

$$a < 4.93 \tag{9.5a}$$

$$r_S > \frac{400320 + 500400a}{0.4928 - 0.1a} \tag{9.5b}$$

The second condition gives a lower bound for r_S when $a = 0$, i.e. $r_S > 8.12 \times 10^5$. Note that this condition cannot be satisfied for non-negative a when $r_S = 3 \times 10^5$, as in the standard parametrisation. Therefore, I will establish a range of values of a and r_S which may lead to spatial pattern formation.

The Jacobian for the spatially homogeneous system (9.1) is

$$J = \begin{pmatrix} -\gamma_S - \frac{\beta b}{b+b_{1/2}} & 0 & -\frac{\beta S b_{1/2}}{(b+b_{1/2})^2} & 0 \\ \frac{\beta b}{b+b_{1/2}} & -(l+a) & \frac{\beta S b_{1/2}}{(b+b_{1/2})^2} & 0 \\ -\frac{\beta b}{b+b_{1/2}} & Nl & -\gamma_b - \frac{\beta S b_{1/2}}{(b+b_{1/2})^2} & 0 \\ 0 & r_c & 0 & -\gamma_c \end{pmatrix} \tag{9.6}$$

Substituting in the values of b and S at $\boldsymbol{u}^*$ gives

$$J = \begin{pmatrix} -\gamma_S - A & 0 & -B & 0 \\ A & -(l+a) & B & 0 \\ -A & Nl & -\gamma_c - B & 0 \\ 0 & r_c & 0 & -\gamma_c \end{pmatrix} \tag{9.7}$$

Where $A = \frac{\beta r_S(Nl-(l+a))-b_{1/2}\gamma_b\gamma_S(l+a)}{r_S(Nl-(l+a))+b_{1/2}\gamma_b(l+a)}$ and $B = \frac{b_{1/2}\gamma_b^2(\beta+\gamma_S)(l+a)^2}{\beta(Nl-(l+a))(r_S(Nl-(l+a))+b_{1/2}\gamma_b(l+a))}$.

The eigenvalues of J establish the stability of the the fixed point at $\boldsymbol{u}^*$ in the spatially homogeneous system; if all eigenvalues have negative real part, the system is stable to perturbations. This establishes a range of values of r_S for each value of a in the range $0 - 4.93$ for which the spatially homogeneous system has a stable fixed point, and thus for which spatial patterning can occur. In order to explore the range of values for which $\boldsymbol{u}$* is stable in the spatially homogeneous system, I used the `eig` function in Matlab to find the largest real part of the eigenvalues of J for combinations of values of a and r_S in the ranges $0 < a < 4.93$ and $0 < r_S < 10^{13}$. This shows that instability only occurs for a very close to the upper limit of 4.93, or for small r_S (Figure 9.1). The boundary between the regions for which the fixed point $\boldsymbol{u}$* in the spatially homogeneous system is stable is complex, but the majority of parametrisations within the region satisfying Conditions 9.5a and 9.5b also lead to a stable fixed point. In general, if $a < 4.7$ and $r_S > 1.21 \times 10^8$ then $\boldsymbol{u}^*$ is stable. This establishes the existence of a persistent disease equilibrium if the system is well-mixed (i.e. there is a stable fixed point for which the bacterial and infected phagocyte populations are non-zero). It is useful to note that the disease-persistent equilibrium does not exist for large a or small r_S; if a is too large, all bacteria are removed from the system, and the persistent infection state is destroyed, while if r_S is insufficiently large, phagocytes are depleted from the system by high levels of cell death, and the infectious steady state cannot persist since the hosts are eliminated too quickly. In both cases, the system would then tend towards the disease-free steady state at $S = r_S/\gamma_S, I = 0, b = 0, c = 0$.

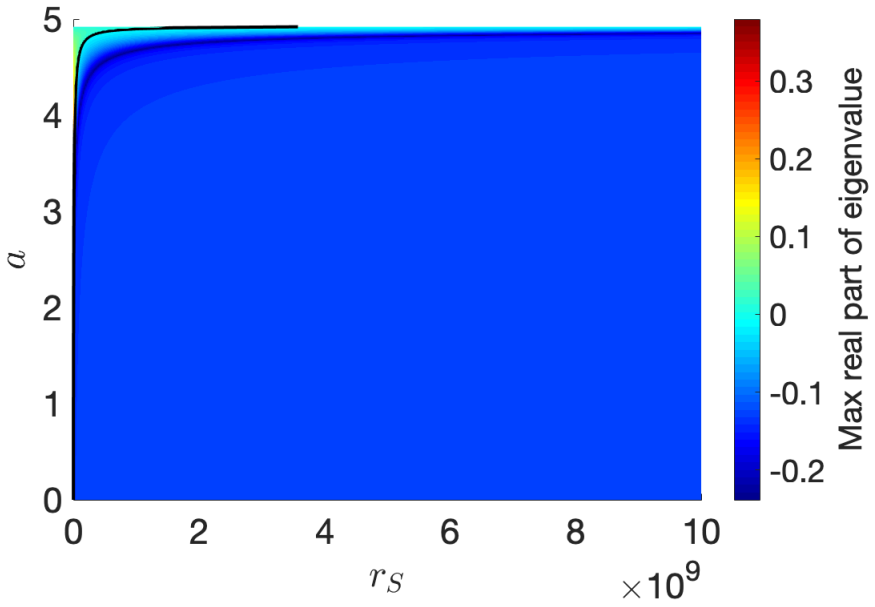


Figure 9.1: Largest real part of eigenvalues of the Jacobian J of the spatially homogeneous system for varying values of the parameters r_S (rate of influx of uninfected phagocytes) and a (apoptosis rate), shown here for $r_S < 10^{10}$. The boundary of the region at which the fixed point exists is shown by a black line; the fixed point exists for all points to the left of and below this line. When all eigenvalues have negative real part, the spatially homogeneous system is stable; this is the case for the majority of values in this plot, with the exception of a very close to the upper limit of 4.93, or r_S very small; the lower limit for r_S is around 1.21×10^8.

9.2.2 Turing instability analysis

We now return to the full PDE system, with transport terms:

$$\begin{aligned}
\frac{\partial S}{\partial t} &= D_S \nabla^2 S - \nabla \cdot (\chi_S S \nabla c) - \beta S \frac{b}{b + b_{1/2}} + r_S - \gamma_S S \\
\frac{\partial I}{\partial t} &= D_I \nabla^2 I - \nabla \cdot (\chi_I I \nabla c) + \beta S \frac{b}{b + b_{1/2}} - (l + a) I \\
\frac{\partial b}{\partial t} &= D_b \nabla^2 b - \beta S \frac{b}{b + b_{1/2}} + N l I - \gamma_b b \\
\frac{\partial c}{\partial t} &= D_c \nabla^2 c + r_c I - \gamma_c c
\end{aligned} \tag{8.1 revisited}$$

We perturb away from the fixed point $\boldsymbol{u}^*$ to assess stability of the system once transport terms are reintroduced. We consider solutions of the following form, corresponding to a plane wave perturbation:

$$\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{u}^* + \hat{\boldsymbol{u}} e^{i \boldsymbol{k} \cdot \boldsymbol{x}} \tag{9.8}$$

We assume $|\hat{\boldsymbol{u}}| \ll 1$.

Substituting Equation 9.8 into Equation 8.1 and linearising gives the following linear ordinary differential equation:

$$\frac{\partial \hat{\boldsymbol{u}}}{\partial t} = J_T \hat{\boldsymbol{u}} \tag{9.9}$$

where J_T is a modified Jacobian including transport terms, evaluated at u^*:

$$J_T = \begin{pmatrix}
-\gamma_S - A - k^2 D_S & 0 & -B & r_S + k^2 S^* \\
A & -(l + a) - k^2 D_I & B & k^2 I^* \\
-A & Nl & -\gamma_c - B - k^2 D_b & 0 \\
0 & r_c & 0 & -\gamma_c - k^2 D_c
\end{pmatrix} \tag{9.10}$$

Here, $k = |\boldsymbol{k}|$. The eigenvalues of J_T at each value of k then determine the stability of the fixed point $\boldsymbol{u}^*$ following spatial perturbation, with the inclusion of transport terms; in particular, if the real parts of all eigenvalues are negative, the components of the perturbation are damped with time, while if any real part of an eigenvalue is positive, the corresponding mode grows with time, and the system moves away from $\boldsymbol{u}^*$. This is a necessary condition for spatial patterning to emerge.

Varying nonspatial parameters

Once again, we investigate the effects of varying the apoptosis rate a, and the rate of influx of uninfected cells r_S, now with transport terms introduced into the system. For all values of a in

the range $0 < a < 4.93$, there are some values of r_S which permit instability of the fixed point $\boldsymbol{u}^*$ in the spatially inhomogeneous system, while $\boldsymbol{u}^*$ is stable to spatial perturbations in the spatially homogeneous system. In particular, for intermediate values of the wavenumber k, the maximum eigenvalue of the Jacobian J_T is positive. In order to establish which parameter combinations of spatial patterning occur, I used the `eig` function in Matlab to determine the largest real part of the eigenvalues of J_T for combinations of parameter sets $0 < a < 4.7$ and $1.21 \times 10^8 < r_S < 10^{13}$. Examples of the maximum real part of the eigenvalues of J_T while varying a, r_S and k, are given in Figures 9.2 and 9.3 respectively; note in particular that for all results shown in these figures, the dispersion relations intercept the x-axis at two points; these give the minimum and maximum values of the wavenumber k for which patterning is possible. The relationship between these maximum and minimum values of k and the parameters a and r_S are shown in Figures 9.4a and 9.4b, while the range of values of k permitting pattern formation are shown in Figure 9.4c.

It is important to note that not all combinations of values of a and r_S can result in spatial patterning; once again, the boundary of the region in which spatial patterning can occur is not simple, but by fitting the edge of the region where patterning occurs in Figure 9.4 using the Matlab function `polyfit` we can see that patterning is possible for most values of a and r_S satisfying $r_S < 1.9356 \times 10^{11} a + 1.6949 \times 10^{11}$. If a is too large or r_S is too small, the system becomes too stable to admit spatial patterning (in both cases, the bacterial population would be eliminated).

The stabilising effect of apoptosis is further confirmed by examining the effect of varying a on the dispersion relations. In particular, increasing a leads to a decreasing overall range of values of k for which patterning is possible; as a increases, the minimum value of k permitting patterning increases, and the maximum value of k decreases (Figures 9.2, 9.4a, 9.4b and 9.4c). This suggests that a wider range of wavelengths of spatial patterns are possible when there is less apoptosis in the system. In contrast, as r_S increases, the range of values of k permitting patterning decreases, with the minimum value of k decreasing (albeit by a very small amount), and the maximum value of k increasing (Figures 9.3, 9.4a, 9.4b and 9.4c). This suggests that a greater influx of uninfected cells into the system would have a destabilising effect on the system, allowing spatial patterning for a larger range of wavelengths.

We can also gain an insight into what sizes of spatial pattern are most likely to occur

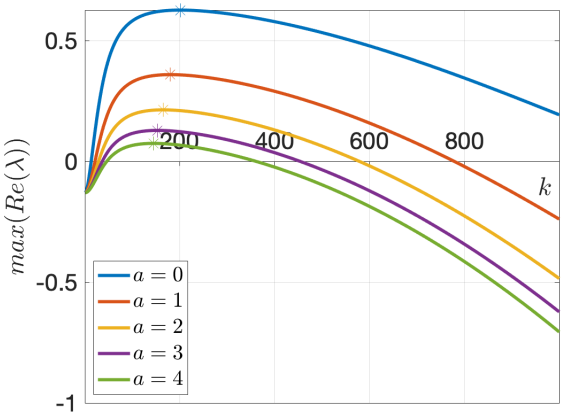


Figure 9.2: Dispersion relations (maximum real part of eigenvalues of the modified Jacobian J_T plotted against values of the perturbation wavenumber k) for varying values of the apoptosis rate a, with standard parametrisation and $r_S = 10^{12}$. For all values of a, the disease persistent fixed point becomes unstable (the maximum real part of eigenvalues of J_T is positive), and spatial patterning can occur, for intermediate values of the wavenumber k. The range of values of k for which instability occurs decreases with increasing a; as a increases, the lower boundary increases and the upper boundary decreases. This shows that apoptosis has a stabilising effect on the system. The dominant mode, marked with a star, also decreases as a increases, suggesting that spatial structures would become larger as a increases.

by considering the wavenumber of the dominant mode (i.e. the value of k which leads to the largest maximum real part of all eigenvalues of J_T). Since we are considering a two-dimensional system, the wavelength of the dominant mode gives an indication of the size of spatial patterning we are likely to see [Murray, 1993]. In particular, increasing a decreases the wavenumber of the dominant mode (which would lead to the formation of larger spatial structures), whereas increasing r_S increases the wavenumber of the dominant mode (leading to smaller spatial structures); a comprehensive view of the dominant wavenumber over varying values of a and r_S is given in Figure 9.4d. In particular, the dominant wavenumbers are all in the range $5-50cm^{-1}$, suggesting wavelengths of $0.2-2mm$. Since we would expect lesion diameters to be closer to the lower end of this range, this suggests that a parameter set with a and r_S near the boundary of stability might be most likely [Richter-Dahlfors et al., 1997, Man et al., 2014a].

Varying motility of infected cells

We now turn our attention to the other process of interest, namely the decrease in motility of infected cells. I have assumed that upon infection, both the diffusion and chemotaxis coefficients are decreased by a constant factor; in particular, $\chi_I = scale \times \chi_S$ and $D_I =$

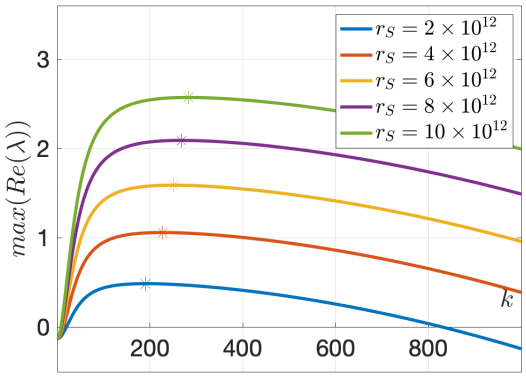


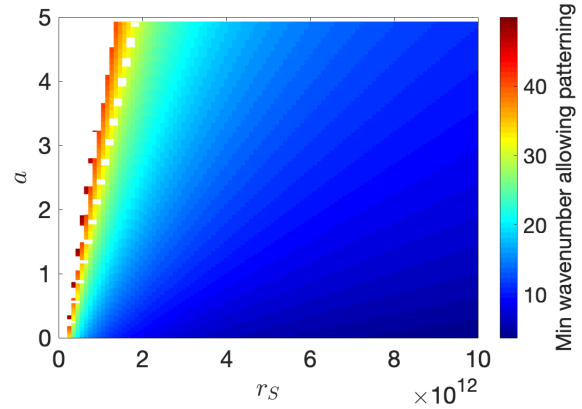
Figure 9.3: Dispersion relations (maximum real part of eigenvalues of the modified Jacobian J_T plotted against values of the perturbation wavenumber k) for varying values of the phagocyte influx rate r_S, with standard parametrisation and $a = 2.5$. For all values of r_S shown here, spatial patterning can occur for intermediate values of the wavenumber k. The range of values of k for which instability occurs increases with r_S; the minimum value of k changes relatively little with varying r_S, but the maximum value of k increases with increasing r_S. This shows that the influx of uninfected phagocytes has a destabilising effect on the system. The dominant mode, marked with a star, also increases as r_S increases, suggesting that the size of spatial structures would decrease with increasing r_S.

$scale \times D_S$ for a motility scale factor ($scale$). I repeated the analysis carried out in the previous section, establishing the maximum real part of the eigenvalues of J_T for various combinations of r_S and a with values of the motility scale between 0 and 1. Examples of dispersion relations are given in Figure 9.5.

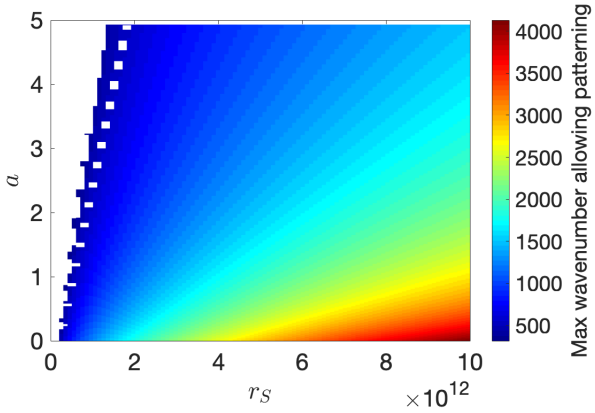
Unlike the results when varying a or r_S, varying the motility of infected cells has little overall effect on the stability of solutions; the minimum, maximum and range of wavenumbers allowing patterning as well as the dominant mode, are broadly unchanged (Figures 9.5, 9.6 and 9.7). Increasing the motility of infected scales slightly decreases the minimum wavenumber allowing patterning, and slightly increases the maximum wavenumber, suggesting that increasing motility of infected cells would lead to a very slight destabilising effect; however, surprisingly, varying the motility of infected cells can have a nonlinear effect on the dominant wavenumber of the system, although increasing motility of infected cells overall tends to increase the dominant wavenumber of the spatial patterns formed.

9.3 Discussion

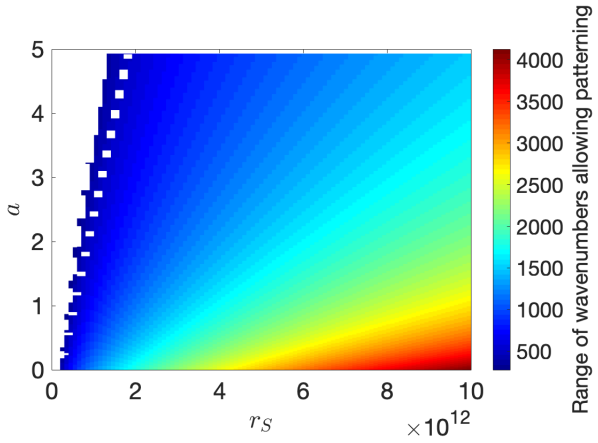
In this chapter, I have analysed the spatial PDE models for lesion formation established in Chapter 8. In particular, I have shown that spatial pattern formation can arise for many



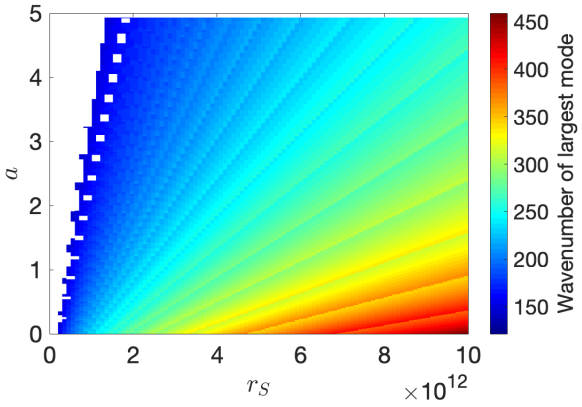
(a) Minimum wavenumber allowing patterning



(b) Maximum wavenumber allowing patterning



(c) Range of wavenumbers allowing patterning



(d) Dominant mode

Figure 9.4: Minimum, maximum and range of values of the perturbation wavenumber k allowing spatial patterning, shown for a range of values of apoptosis rate a and phagocyte influx rate r_S (standard values are used for other parameters). Parameter values in white areas do not display spatial patterning, since there is no instability upon introduction of transport terms. For fixed values of r_S, increasing a increases the minimum value of k allowing patterning, and decreases the maximum value of k allowing patterning. The viable range of values of k increases as r_S increases, and decreases as a increases. The wavenumber of the dominant mode generally decreases with increasing a, and increases with increasing r_S. Overall, this suggests apoptosis has a stabiising effect on the system, while phagocyte influx has a destabilising effect.

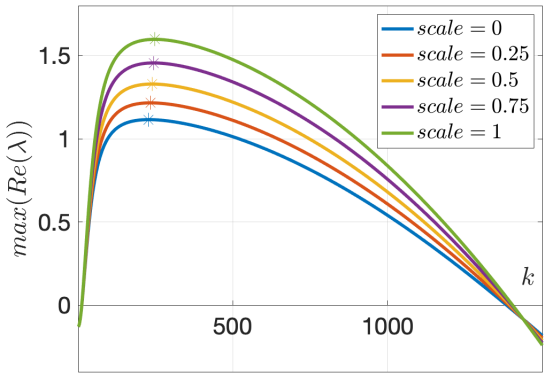


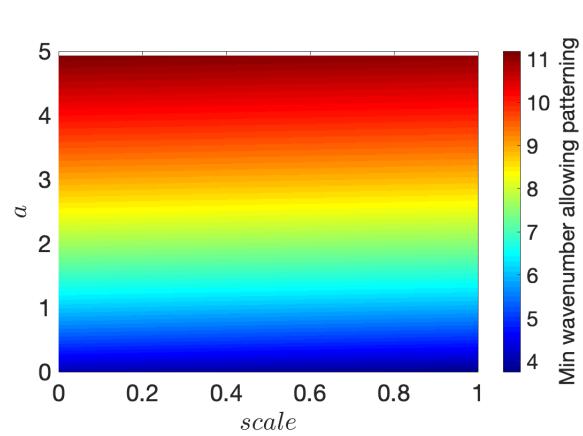
Figure 9.5: Dispersion relations (maximum real part of eigenvalues of the modified Jacobian J_T plotted against values of the perturbation wavenumber k) for varying values of a scaling parameter linking rate of movement of infected and uninfected phagocytes, with standard parametrisation and $a = 2.5, r_S = 5 \times 10^{12}$. The minimum wavenumber allowing patterning decreases slightly as motility of infected cells increases, and the maximum wavenumber increases slightly. However, the relative change in minimum and maximum wavenumber are very small; the dominant mode is also changed very little. This suggests that decreasing the motility of infected cells has a very slight stabilising effect.

different parametrisations of the model, and I have derived necessary conditions for spatial pattern formation on the parameters.

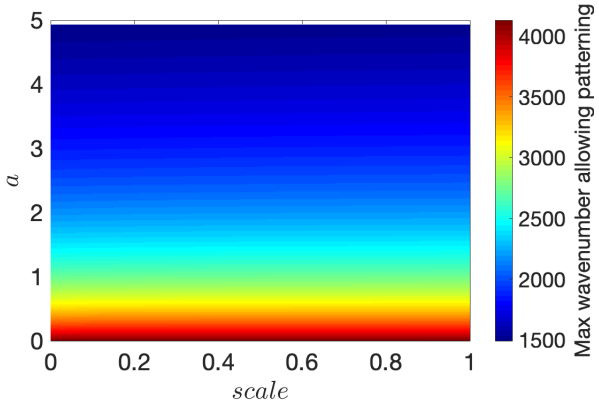
Notably, there is a relationship between apoptosis rate and the rate of input of uninfected phagocytes into the system, which establishes a minimum value for the phagocyte input rate as a function of the apoptosis rate. This suggests that a high rate of uninfected phagocyte input may be especially important in the development of lesions, and this process cannot be excluded from the model. I have also demonstrated that spatial patterning is not possible if apoptosis rates are too high; this suggests that apoptosis may limit the ability of the system to form lesions, since it has a stabilising effect on the system. I have also shown that, surprisingly, changing the comparative motility rates in infected and uninfected cells does not have a particularly large effect on the ability of the system to form spatial patterns, although the dominant mode of the resultant spatial pattern is affected. Finally, I have established the wavenumber of the dominant mode for the parameter sets considered here, which gives an indication of the size of the resulting spatial structures.

9.3.1 Further work

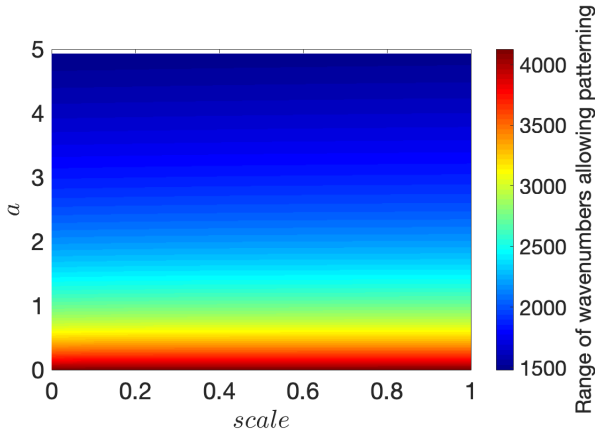
It would naturally be possible to extend this analysis to consider the stability behaviour when varying any other parameter in the model, or even to consider a larger-dimensional space



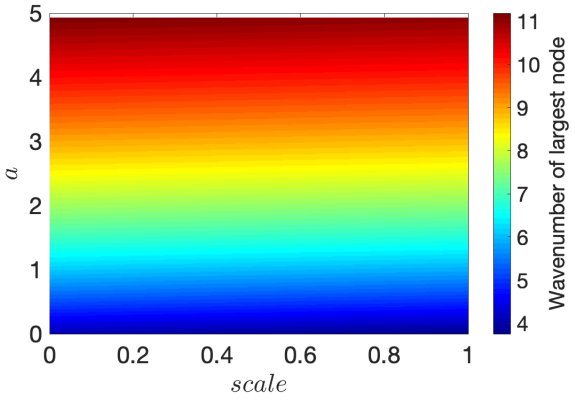
(a) Minimum wavenumber allowing patterning



(b) Maximum wavenumber allowing patterning

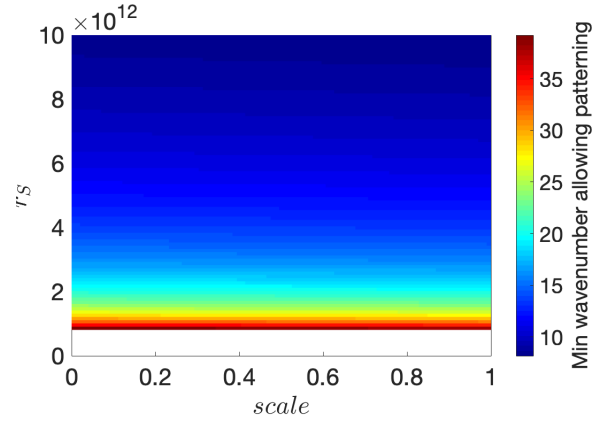


(c) Range of wavenumbers allowing patterning

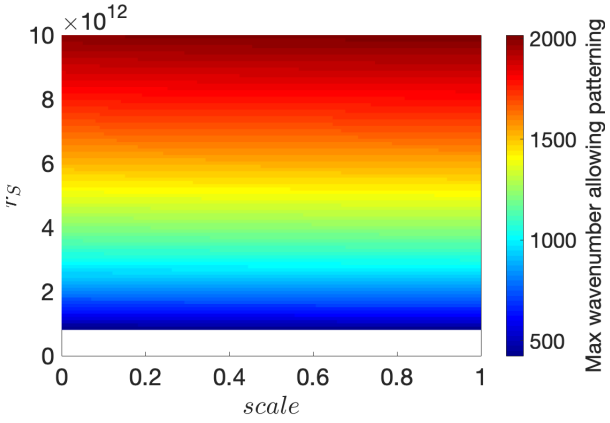


(d) Dominant mode

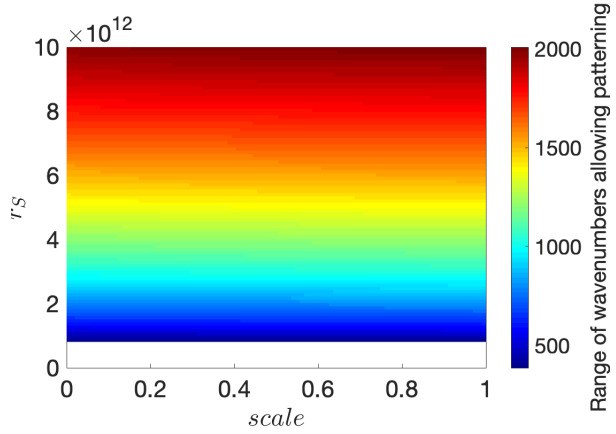
Figure 9.6: Minimum, maximum and range of values of the perturbation wavenumber k allowing spatial patterning, shown for a range of values of the motility scale for infected cells and apoptosis rate a, with $r_S = 5 \times 10^{12}$ (standard values are used for other parameters). Increasing the motility of infected cells slightly decreases the minimum wavenumber allowing patterning, and slightly increases the maximum wavenumber allowing patterning as well as the range of wavenumbers allowing patterning. Increasing the motility of cells also overall tends to increase the wavenumber of the dominant mode. This suggests that increasing the motility of infected cells slightly decreases the stability of the system, and increasing motility of cells may lead to larger wavenumbers in the spatial patterns formed.



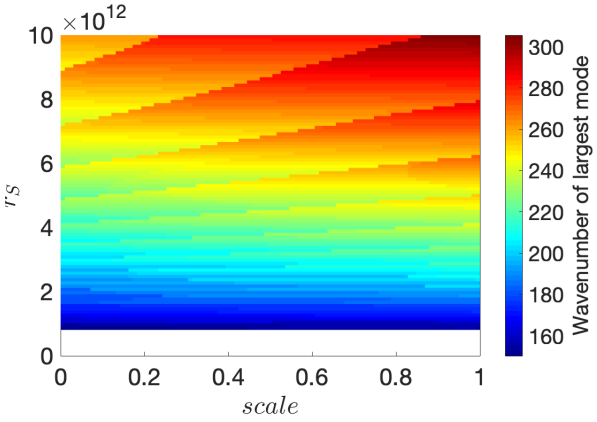
(a) Minimum wavenumber allowing patterning



(b) Maximum wavenumber allowing patterning



(c) Range of wavenumbers allowing patterning



(d) Dominant mode

Figure 9.7: Minimum, maximum and range of values of the perturbation wavenumber k allowing spatial patterning, shown for a range of values of the motility scale for infected cells and the phagocyte influx rate r_S, with $a = 5 \times 10^{12}$ (standard values are used for other parameters). Similarly to the results for varying a (Figure 9.6), increasing the motility of infected cells slightly decreases the minimum wavenumber allowing patterning, and slightly increases the maximum wavenumber allowing patterning as well as the range of wavenumbers allowing patterning. Once again, overall increasing the motility of cells tends to increase the wavenumber of the dominant mode. This suggests that increasing the motility of infected cells slightly decreases the stability of the system, and may lead to larger wavenumbers in the spatial pattern formed.

in which stability can occur by varying multiple parameters at once; however, this analysis quickly becomes extremely unwieldy and difficult to interpret, so I have confined my interest to the parameters which are most biologically interesting or for which a reliable estimate is not available.

Furthermore, the wavenumbers of the dominant modes give an indication of the size of the spatial patterns which can develop in these systems; these may be directly comparable to the diameter of lesions *in vivo*. It would be extremely useful to develop an experimental dataset including the measured size of lesions in histological samples; this would allow us to establish further conditions on the parameters varied in this chapter, by demonstrating which parameters give rise to spatial pattern structures of a similar size to experimentally observed lesions.

While this analysis has enabled us to establish some bounds on some of the parameters in the model if lesions are the result of Turing instability-like spatial pattern formation, and also to gain a clearer insight into which processes might promote or oppose spatial pattern formation, there are limits to what such an approach allows us to establish. While we can find the fastest-growing wavenumbers of the spatial patterns which may form, it is still unclear what sort of spatial patterns are possible in this system. Furthermore, this analysis focuses on long-term behaviour of the system rather than demonstrating how the spatial patterns actually form. In the next chapter, I will therefore present numerical solutions for the PDE spatial model, building on the insights obtained from the analytical approach presented above.

Chapter 10

Simulations of the spatial model of lesion formation

10.1 Introduction

As demonstrated in the previous chapter, stable spatial pattern formation is possible for some parametrisations of the lesion formation model; this could correspond to the distinctive spatial arrangement of lesions. However, the analytical work carried out so far can only give us limited insight into what kinds of patterns are possible in biologically realistic scenarios, and whether these correspond to *in vivo* observations of lesion formation. This motivates further investigation of the possible solutions of these models; since they are not analytically solvable, I will use numerical integration techniques to approximate solutions for different parametrisations of the model over space and time.

In this chapter, I will use the insight into stability arising from different parametrisations from the previous chapter to guide an analysis of the lesion formation models using simulated solutions. I will explain the methodology used for solving the systems of PDEs in the models, then present solutions for a variety of parametrisations. In particular, I will demonstrate that isolated areas of bacteria input can give rise to round, lesion-like areas of concentrated infected phagocytes, which are stable in size for the first 1-2 weeks of infection, although at later times these lesions begin to grow in radius at a linear rate. Similarly, infected phagocyte and extracellular bacteria levels are initially very low, but begin to grow at an approximately quadratic rate at around 1-2 weeks after the initial infection.

Thus, at least at early times, the solutions for this model coincide with the account of

lesion growth laid out in Chapter 8. It is likely that other biological processes may intervene, either during the initial phase of constant lesion size, or early in the growth phase, to prevent unbounded linear growth. I also establish that increasing the apoptosis rate and decreasing the phagocyte influx rate may lead to lesions which are stable in size for longer, and grow more slowly during the growth phase. Meanwhile, the level of motility of infected phagocytes determines the initial lesion size, as well as the rate of growth during the growth phase. This suggests that apoptosis and reduction of infected cell motility both play a key role in stabilising the spatial spread of infection, as well as the bacterial burden, while the influx of phagocytes must be carefully regulated to prevent infection growth.

10.2 Simulation methods

Once again, we will be considering the spatial PDE model originally presented in Chapter 8:

$$\begin{aligned}
\frac{\partial S}{\partial t} &= D_S \nabla^2 S - \nabla \cdot (\chi_S S \nabla c) - \beta S \frac{b}{b + b_{1/2}} + r_S - \gamma_S S \\
\frac{\partial I}{\partial t} &= D_I \nabla^2 I - \nabla \cdot (\chi_I I \nabla c) + \beta S \frac{b}{b + b_{1/2}} - (l + a) I \\
\frac{\partial b}{\partial t} &= D_b \nabla^2 b - \beta S \frac{b}{b + b_{1/2}} + N l I - \gamma_b b \\
\frac{\partial c}{\partial t} &= D_c \nabla^2 c + r_c I - \gamma_c c
\end{aligned} \tag{8.1 revisited}$$

With the exception of the apoptosis rate a, the rate of influx of uninfected phagocytes r_S, and the rates of motility of infected cells (the diffusion coefficient D_I and chemotaxis rate χ_I), the standard parametrisation established in Chapter 8 is used (Table 10.1).

For simplicity, I set the domain to be a two-dimensional square $[0, L]^2$; while naturally the tissue in which lesions are formed is three-dimensional, an initial model in two dimensions simplifies the problem while still allowing spatial patterning in the solutions.

To solve the PDEs in System 8.1, I used the method of lines; this is a standard approach to numerically solving PDEs by discretising the domain on which they are to be solved, and approximating the solutions at each point by an ODE system. I discretised the domain using a $n \times n$ grid, with gridlines spaced a distance L/n apart. The concentrations of reactants are calculated at each point in the grid; in particular, variable X is discretised as $X_{i,j}(t) = X((\frac{iL}{n}, \frac{jL}{n}), t)$ for $i, j \in \{0, 1, \cdots, L\}$. The PDE system can then be converted to a system of ODEs, which can be solved using a standard numerical method. For example, the Laplace

Parameter	Value used	Unit
D_S	8.64×10^{-7}	$cm^2 d^{-1}$
D_b	2.01×10^{-5}	$cm^2 d^{-1}$
D_c	3×10^{-2}	$cm^2 d^{-1}$
χ_S	1.61×10^{-2}	$cm^4 g^{-1} d^{-1}$
β	0.1	d^{-1}
$b_{1/2}$	6×10^5	cm^{-2}
γ_S	0.03	d^{-1}
γ_I	0	d^{-1}
l	0.8	d^{-1}
N	7.16	NA
γ_b	27.8	d^{-1}
r_c	2.86×10^{-12}	d^{-1}
γ_c	55.45	d^{-1}

Table 10.1: Parameters used in simulations of the spatial PDE model of lesion formation

operator applied to variable X can be approximated at gridpoint (i, j) is as follows:

$$(\nabla^2 X)_{i,j} \approx \frac{n^2}{L^2}(X_{i+1,j} + X_{i-1,j} + X_{i,j+1} + X_{i,j-1} - 4X_{i,j}) \tag{10.1}$$

In order to define the reflecting Neumann boundary conditions, we can define the values of X at fictitious grid points $(-1, j), (n + 1, j), (i, -1), (i, n + 1)$ as follows:

$$X_{-1,j} = X_{1,j} \tag{10.2a}$$
$$X_{L+1,j} = X_{L-1,j} \tag{10.2b}$$
$$X_{i,-1} = X_{i,1} \tag{10.2c}$$
$$X_{i,L+1} X_{i,L-1} \tag{10.2d}$$

This discretisation is mostly straightforward except for the chemotaxis term. It would be possible to apply a similar scheme to the Laplace operator terms (Equation 10.1); however, there are often difficulties with smoothness and positivity of solutions in solving chemotaxis-type terms in PDE systems. In fact, applying a straightforward discretisation scheme to the chemotaxis terms in this particular system gives rise to negative solutions. This is a well-documented issue with solving reaction-diffusion-chemotaxis systems numerically [Gerisch et al., 2001, Gerisch and Verwer, 2002, Gerisch and Chaplain, 2006], particularly when the chemotaxis parameter is many orders of magnitude larger than diffusion coefficient (as is the

case with this model). The discontinuities often arise as the system displays concentrated areas with large abundances, which are difficult to render in a discretised system. Furthermore, the reaction-diffusion terms often lead to a stiff system, which requires a specialised approach for solving; coupled with the relatively large terms due to chemotaxis, this often results in very small timesteps, ultimately causing the method to be extremely computationally intensive.

Here I have chosen to use a method based on an approach designed by Gerisch *et al.* [Gerisch et al., 2001]. In particular, this method uses spatial discretisation coupled with a Van Leer flux limiter [Sweby, 1984], which ensures the positivity of solutions for all time. This method also uses a splitting scheme to solve the chemotaxis and reaction-diffusion parts of the equation separately; thus the stiff parts of the PDE can be treated separately using an appropriate solver, while the solution to the chemotaxis part of the system can be derived using larger timesteps. The method is broadly as outlined in a previous study [Gerisch et al., 2001], although I have adapted the method so that the flux is two-dimensional.

The diffusion parts of the equations were discretised using the system outlined above (10.1). The chemotaxis parts were approximated using the following scheme:

$$\nabla \cdot (\chi X \nabla c)_{i,j} = F^x_{i,j} + F^y_{i,j} \tag{10.3}$$

where $F^x_{i,j}$ and $F^y_{i,j}$ are the fluxes due to chemotaxis in the x and y direction respectively. These fluxes are functions of the velocities ($f^x_{i,j}$ and $f^y_{i,j}$) of the reactant X in the x and y directions respectively, which are approximated as follows:

$$f^x_{i,j} = \frac{X_{i,j}}{2h}(c_{i+1,j} - c_{i-1,j}) \tag{10.4a}$$

$$f^y_{i,j} = \frac{X_{i,j}}{2h}(c_{i,j+1} - c_{i,j-1}) \tag{10.4b}$$

The flux in the x-direction is approximated thus:

$$F^x_{i,j} = \begin{cases} -\frac{\chi}{h}\left[\left(1 + \frac{1}{2}\Phi(r_{i+1/2,j})\right) - \frac{\Phi(r_{i-1/2,j})}{2r_{i-1/2,j}}\right](f^x_{i,j} - f^x_{i-1,j}) & \text{if } (X_{i,j} > 0 \text{ and } f_{i,j} \geq 0) \\ & \text{or } (X_{i,j} = 0 \text{ and } f_{i-1,j} \geq 0) \\ -\frac{\chi}{h}\left[\left(1 + \frac{1}{2}\Phi(r^{-1}_{i+1/2,j})\right) - \frac{\Phi(r^{-1}_{i+3/2,j})}{2r^{-1}_{i+3/2,j}}\right](f^x_{i+1,j} - f^x_{i,j}) & \text{if } (X_{i,j} > 0 \text{ and } f_{i,j} < 0) \\ & \text{or } (X_{i,j} = 0 \text{ and } f_{i+1,j} \leq 0) \\ 0 & \text{otherwise} \end{cases} \tag{10.5a}$$

$$\Phi(r) = \frac{|r|+r}{1+|r|} \tag{10.5b}$$

$$r_{i+1/2,j} = \frac{f^x_{i+1,j} - f^x_{i,j}}{f^x_{i,j} - f^x_{i-1,j}} \tag{10.5c}$$

Here the function $\Phi(r)$ is the van Leer limiter function, which ensures non-negativity of $X_{i,j}$.

The flux in the y direction is approximated in an analogous fashion:

$$F^y_{i,j} = \begin{cases} -\frac{\chi}{h}\left[\left(1+\frac{1}{2}\Phi(r_{i,j+1/2})\right) - \frac{\Phi(r_{i,j-1/2})}{2r_{i,j-1/2}}\right](f^y_{i,j} - f^y_{i,j-1}) & \text{if } (X_{i,j} > 0 \text{ and } f_{i,j} \geq 0) \\ & \text{or } (X_{i,j} = 0 \text{ and } f_{i,j-1} \geq 0) \\ -\frac{\chi}{h}\left[\left(1+\frac{1}{2}\Phi(r^{-1}_{i,j+1/2})\right) - \frac{\Phi(r^{-1}_{i,j+3/2})}{2r^{-1}_{i,j+3/2}}\right](f^y_{i,j+1} - f^y_{i,j}) & \text{if } (X_{i,j} > 0 \text{ and } f_{i,j} < 0) \\ & \text{or } (X_{i,j} = 0 \text{ and } f_{i,j+1} \leq 0) \\ 0 & \text{otherwise} \end{cases} \tag{10.6a}$$

$$r_{i,j+1/2} = \frac{f^y_{i,j+1} - f^y_{i,j}}{f^y_{i,j} - f^y_{i,j-1}} \tag{10.6b}$$

Together with the reaction and diffusion parts of the system, this gives a system of ODEs of the following form:

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = f_1(\boldsymbol{u}) + f_2(\boldsymbol{u}) \tag{10.7}$$

where f_1 represents the chemotaxis terms and f_2 represents the reaction and diffusion terms. Thus we have separated the ODEs into stiff (f_2) and non-stiff (f_1) parts. We can now solve Equation 10.7 using an iterative process as follows:

Suppose an estimate $\boldsymbol{u}^*$ for the solution of System 10.7 is known at time t^*. We define a series of variables $\boldsymbol{z}_1(t), \boldsymbol{z}_2(t), \boldsymbol{z}_3(t)$ as the solutions to the following equations:

$$\begin{aligned} \tfrac{\mathrm{d}\boldsymbol{z}_1}{\mathrm{d}t} &= f_1(\boldsymbol{z}_1) \quad \boldsymbol{z}_1(t^*) = \boldsymbol{u}^* \\ \tfrac{\mathrm{d}\boldsymbol{z}_2}{\mathrm{d}t} &= f_2(\boldsymbol{z}_2) \quad \boldsymbol{z}_2(t^*) = \boldsymbol{z}_1(t^* + \tau/2) \\ \tfrac{\mathrm{d}\boldsymbol{z}_3}{\mathrm{d}t} &= f_1(\boldsymbol{z}_3) \quad \boldsymbol{z}_3(t^* + \tau/2) = \boldsymbol{z}_2(t^* + \tau) \end{aligned} \tag{10.8}$$

The calculated solution for $\boldsymbol{z}_3(t^* + \tau)$ can then be used as an approximation for $\boldsymbol{u}(t^* + \tau)$. Note that since only $\boldsymbol{z}_2$ relies on integrating f_2, this step is the only part of the problem which requires solving the stiff system resulting from reaction-diffusion terms, whereas deriving $\boldsymbol{z}_1$ and $\boldsymbol{z}_3$ relies only on the chemotaxis terms. This scheme can be applied iteratively, using

the ICs for the model to give the initial values of $\boldsymbol{u}^*$, using a stiff ODE solver to solve the equation for z_2, and a non-stiff ODE solver to solve the equations for z_1 and z_3.

I used this technique for a 100×100 grid in Matlab, using a simple fourth-order Runge-Kutta approach to solve the equations in System 10.8 involving f_1, and the `ODE15s` function to integrate the equations involving f_2. For the latter case, the system is very large, so it is necessary to use a sparse Jacobian both for speed and maximum memory considerations. Overall, this method was successful at keeping solutions positive but still efficient. I ran simulations for various parameter values and initial conditions from 0 to 30 days - lesions should be well established by around 1 week following infection, but it was also useful to examine the longer-term behaviour of the system.

Once I had generated a simulation for a particular parameter set, I isolated the size of the lesions at each timepoint by converting the infected phagocyte concentrations into a binary image representation using the Matlab function `imbinarize`; this isolated areas of high concentration, which were deemed to lie within a lesion. I also tracked the total concentrations of infected phagocytes and extracellular bacteria over time, in order to analyse the growth of infection. This was accomplished simply by integrating the respective variables over the spatial domain at each timepoint.

10.3 Results

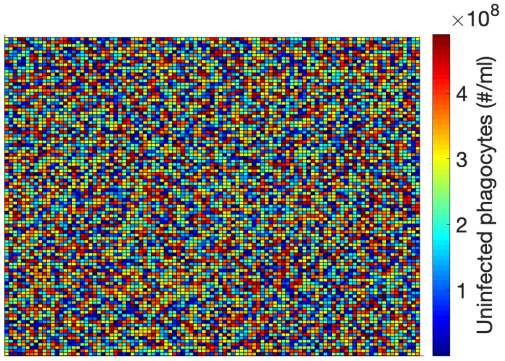
10.3.1 Results for basic parameter set

I initially ran simulations for a $1mm \times 1mm$ square, with an initial bacterial density of zero everywhere except at the centre of the grid, with one square set to a density of $200cm^{-2}$ (Figure 10.1). Over time, a concentrated area of infected phagocytes and bacteria appears in the centre of the domain, around the point at which bacteria are initially introduced (Figure 10.2). Uninfected phagocyte concentration in this central area is comparatively very low, although chemokines are also concentrated in this area, and there is also a diffuse area of lower levels of chemokine radiating outward from the centre of the area of highest concentration. This defined area of phagocytes and bacteria corresponds well to the definition of a lesion outlined previously (note that the area of infected phagocytes is slightly larger than the area of bacteria, suggesting that the bacteria are confined by infected phagocytes). We only see one lesion-like area forming, with no seeding of further lesions across the domain.

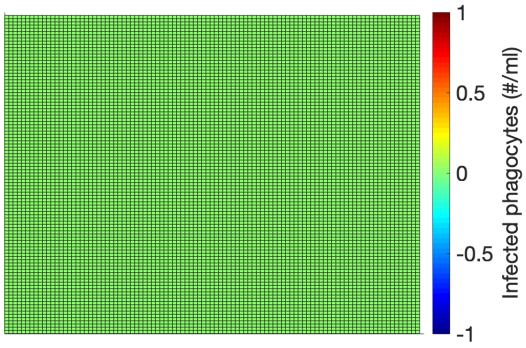
The size of the lesion is approximately constant at early times (until around 1-2 weeks post-infection). Thereafter, the lesion begins to grow at an approximately linear rate over time (Figures 10.3 and 10.4). I will refer to these distinct phases in the infection as the 'plateau' and 'growth' phases. The concentration of bacteria and infected phagocytes at the centre of the lesion becomes very high at later times, increasing exponentially to the point that the stiff ODE solver is unable to execute a timestep within an accuracy tolerance of 3×10^{14}. Thus, the reaction system becomes too stiff even for a robust stiff ODE solver such as `ode15s`. Up to the point that the ODE solver fails, we also see the total concentration of infected phagocytes and bacteria increasing over the first thirty days of infection; after an initial period during which time both concentrations are very low, both begin to increase approximately quadratically (Figure 10.5). There are therefore 'plateau' and 'growth' phases for the total concentrations of bacteria and infected phagocytes, analogous to those observed for the lesion size.

These results suggest that with the standard parametrisation, the model quickly reaches a reasonably stable state which appears similar to experimental accounts of lesion formation. However, at least for this domain and the bacterial initial conditions tested here, this spatial structure is ultimately unstable. This suggests that we do not, in fact, see stable spatial patterns as suggested by the Turing instability analysis of the previous chapter. The infected phagocytes and bacteria are spatially arranged much as we would expect for lesions during the initial plateau period, when the lesion size is constant, although they are somwhat smaller than experimental observations (around 25-100 μm [Richter-Dahlfors et al., 1997, Man et al., 2014a]). However, following this plateau, the lesion-like areas grow linearly over time, rather than remaining limited in size and seeding further lesions elsewhere. This late-stage growth has not been observed experimentally, but it is worth noting that this growth phase tends to occur fairly late during the infection, and it is possible that another process intervenes to stabilise the growing lesions and prevent further growth (for example, the introduction of fibroblasts, which have not been included in this model, or the intervention of adaptive immune responses at later times).

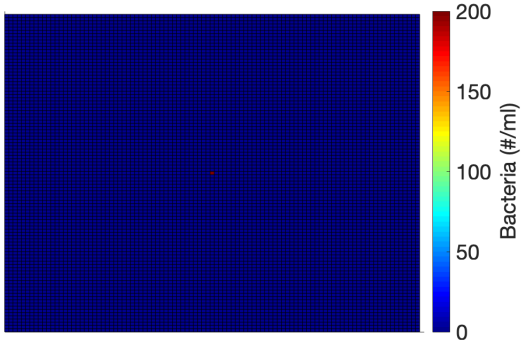
In order to be able to compare simulation results easily between different parametrisations, and to be able to divide the solutions into plateau and growth phases, I fitted a bilinear model to the calculated lesion radius over time (Figure 10.4). This was very similar to the bilinear fit carried out in Chapter 6 on the traces extracted from the time series showing in-



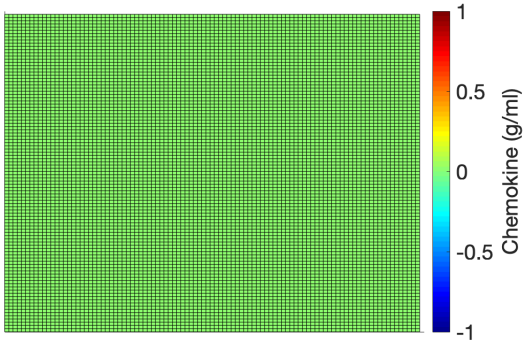
(a) Uninfected phagocytes



(b) Infected phagocytes

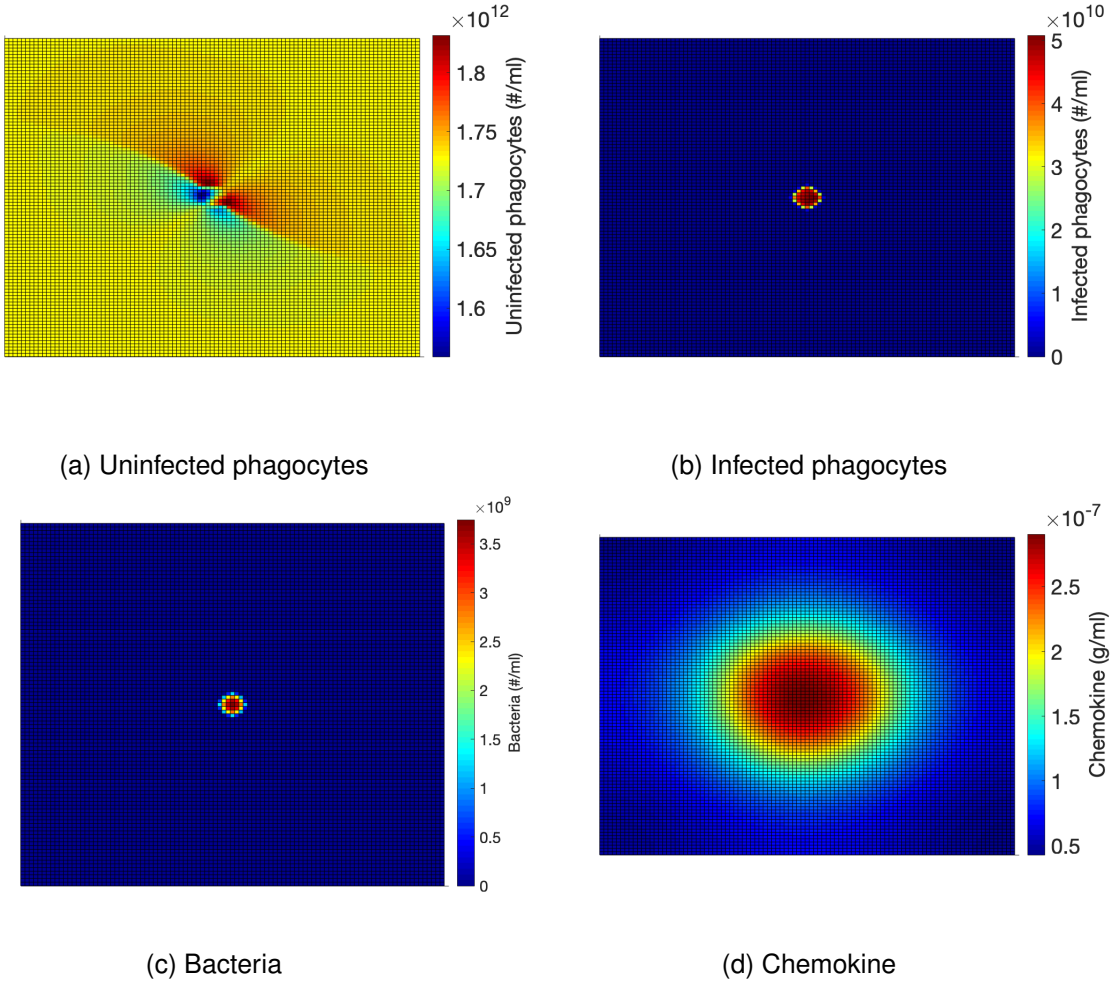


(c) Bacteria



(d) Chemokine

Figure 10.1: Initial conditions for simulations of the spatial PDE lesions model. Uninfected phagocytes are initially uniform randomly distributed with mean $2.45 \times 10^8 cm^{-2}$. Bacteria concentration is zero everywhere except for one cell in the centre of the domain with concentration $200 cm^{-2}$. Infected phagocyte and chemokine concentrations are zero throughout the domain.



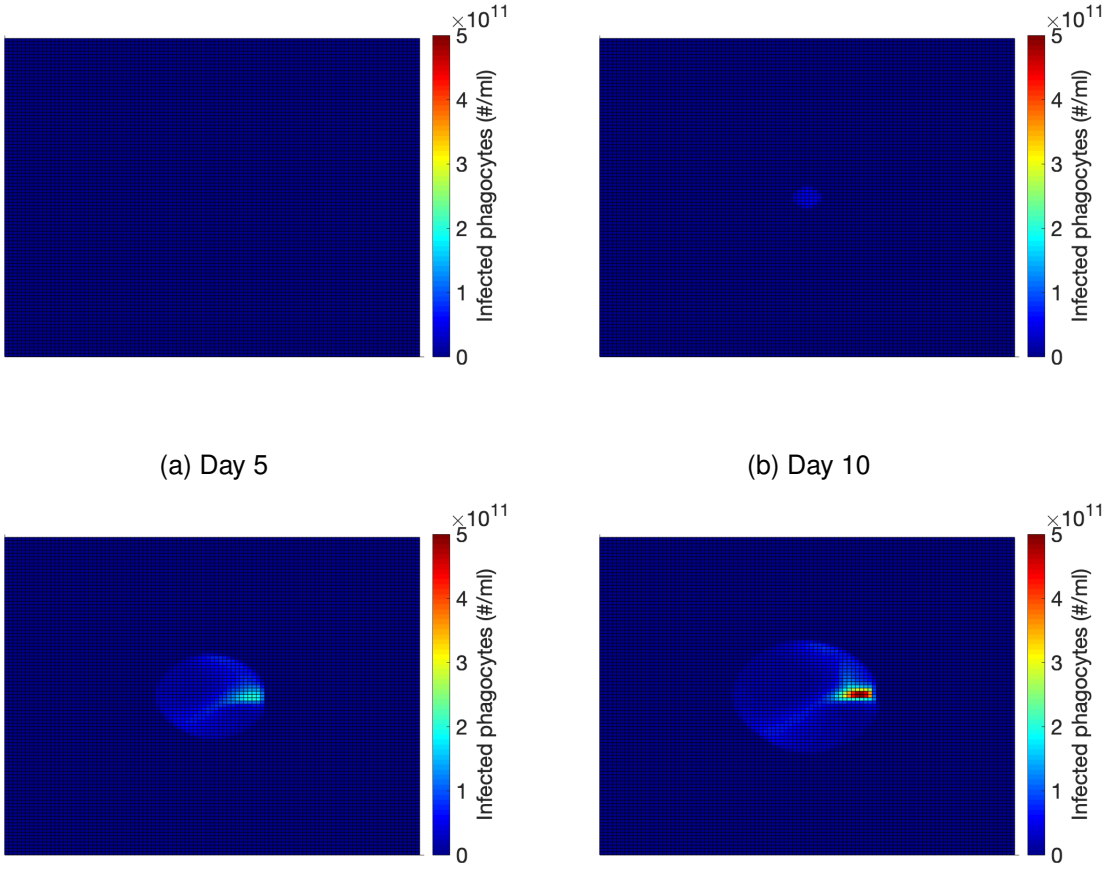
(a) Uninfected phagocytes

(b) Infected phagocytes

(c) Bacteria

(d) Chemokine

Figure 10.2: Results from lesion formation calculated numerically, with apoptosis rate $a = 2.5$, phagocyte influx rate $r_S = 2 \times 10^{11}$ and no movement of infected phagocytes, 10 days from start of infection. Bacteria and infected phagocytes are confined in a roughly circular area, from which uninfected phagocytes are largely absent; chemokines are also concentrated in this area, but diffuse outwards. This area corresponds to a lesion.



(a) Day 5

(b) Day 10

(c) Day 15

(d) Day 17

Figure 10.3: Concentrations of infected phagocytes between 0 and 17 days post-infection calculated numerically, with apoptosis rate $a = 2.5$, phagocyte influx rate $r_S = 2 \times 10^{11}$ and no movement of infected phagocytes. The area of infected phagocytes is centred around the input of bacteria and is roughly circular, and initially very small, although growth begins to occur at around day 8 post-infection.

flammasome growth, although I assumed the lesion radius is constant for the initial phase. The model is as follows:

$$r_i = \beta_0 + \beta_1(t_i - t_b)^+ + \epsilon_i \tag{10.9}$$

where r_i is the ith measurement of the lesion radius, t_i is the ith timepoint, t_b is the breakpoint, and ϵ_i is a normally-distributed random error for the ith measurement.

Once again, $(t_i - t_{b_j})^+$ is defined as follows:

$$(t_i - t_{b_j})^+ = \begin{cases} t_i - t_{b_j} & \text{if } t_i > t_{b_j} \\ 0 & \text{otherwise} \end{cases} \tag{2.2 revisited}$$

β_0 then corresponds to the size of the lesion during the plateau phase, the breakpoint t_b marks the end of the plateau phase and the start of the growth phase, and β_1 represents the growth rate during the growth phase.

I carried out a similar fit for the total concentration of infected phagocytes and bacteria; in this case, however, I assumed that the concentration during the first section is zero, and increases quadratically in the second section (this gives a closer match to results than bilinear or exponential fits). The model is as follows:

$$I_i^{tot} = \beta_I(t_i - t_b^I)^+ + \epsilon_i^I \tag{10.10a}$$
$$b_i^{tot} = \beta_b(t_i - t_b^b)^+ + \epsilon_i^b \tag{10.10b}$$

where I_i^{tot} and b_i^{tot} are the ith measurement of the total infected phagocyte and bacteria concentrations respectively, t_b^I and t_b^b are the calculated breaktimes for infected phagocytes and bacteria, and ϵ_i^I and ϵ_i^b are the errors for the ith measurement of infected phagocytes and bacteria. In this case, β_I and β_b are the growth rates during the growth phase, and t_b^I and t_b^b are breakpoints between the plateau and growth phases. Note that I do not assume that the breakpoints t_b, t_b^I and t_b^b are equal.

Examples of the fits are given in Figures 10.4 and 10.5.

10.3.2 Varying apoptosis rate and phagocyte input rate

In order to discern the effect of varying the apoptosis rate a and the phagocyte input rate r_S, I fitted the bilinear models described above to the lesion area, the total concentration of infected phagocytes, and the total concentration of bacteria over time for simulated results for a range of values of a and r_S, with no movement of infected phagocytes. I determined the

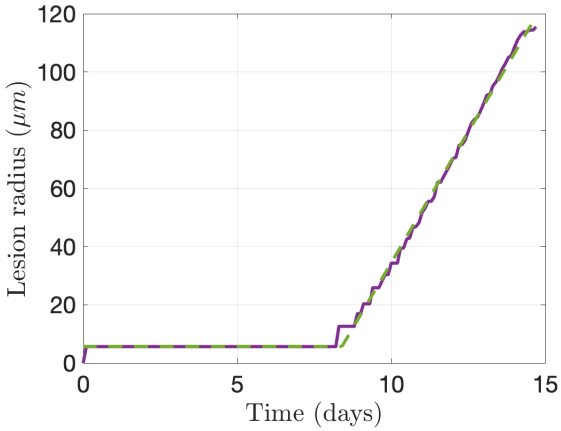
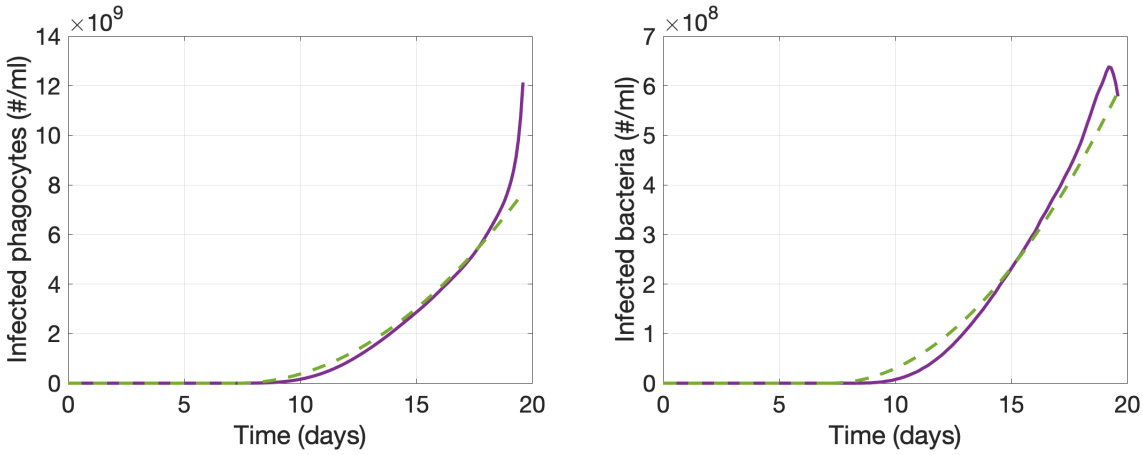


Figure 10.4: Radius of lesion between 0 and 15 days post-infection (shown in purple), extracted from numerical solutions of the PDE model, with apoptosis rate $a = 2.5$, phagocyte influx rate $r_S = 2 \times 10^{11}$ and no movement of infected phagocytes. After the initial plateau phase, the lesion grows over time at an approximately linear rate (growth phase). A fitted bilinear model with constant first section is shown in green.



(a) Infected phagocytes

(b) Bacteria

Figure 10.5: Total concentration of infected phagocytes and bacteria between 0 and 19 days post-infection (shown in purple), extracted from numerical solutions of the PDE model, for apoptosis rate $a = 2.5$, phaogcyte influx rate $r_S = 2 \times 10^{11}$ and no movement of infected phagocytes. Following an initial plateau phase during which the number of infected phagocytes and bacteria is approximately zero, the populations of both infected phagocytes and bacteria begin to grow at an approximately quadratic rate (growth phase). A bilinear fit with first section zero and second section quadratic is shown in green.

range of values of a and r_S to be used from the Turing analysis carried out in the previous chapter. The results are shown in Figures 10.6 and 10.7.

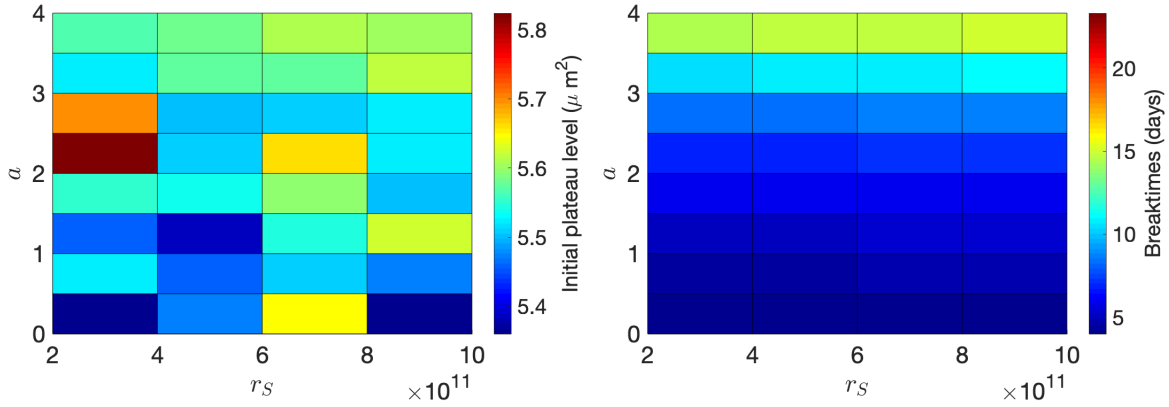
The lesion size throughout the infection is sensitive to changes in both a and r_S (Figure 10.6). While there is no clear relationship between a or r_S and the size of the lesion during the initial plateau, the length of the plateau phase is certainly dependent on a and r_S; for larger values of a and smaller values of r_S, the time spent in the early plateau phase is longer. This suggests that higher rates of apoptosis and lower rates of phagocyte influx lead to lesions which remain at a small, stable size for longer before linear growth begins. The rate of this linear growth increases as a decreases and r_S increases, suggesting again that for higher apoptosis rates and lower phagocyte influx rates, the eventual rate of lesion growth is slower. Likewise, the maximum time reached by the solver is higher for higher a and lower r_S since a slower rate of lesion growth is likely to result in a longer time before the solver fails.

The growth of infection is also sensitive to a; as a increases, the fitted breaktimes (i.e., the end of the initial plateau) for both infected phagocytes and bacteria increase. There is no clear relationship between the length of the plateau phase and r_S. This suggests that for higher apoptosis rates the overall levels of infection remain low for longer. On the other hand, the rate of the final increase decreases as a increases, and once again there is no clear relationship with r_S. This suggests that when the lesions do begin to increase in size, this occurs at a slower rate when apoptosis rates are higher.

Similar results are observed for the total concentrations of bacteria and infected phagocytes (Figure 10.7); the breaktimes for both variables increase as a increases, and the final growth rate decreases as a increases. This suggests that for higher rates of apoptosis, the infection level remains stable for longer, and the growth of the infection during the growth phase is slower.

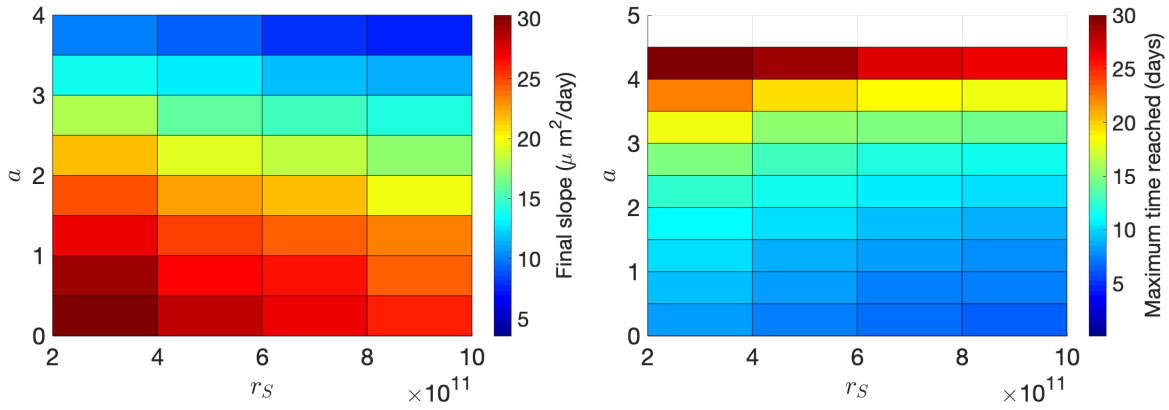
Note that the breaktimes for the bacterial and infected phagocyte concentrations are generally earlier than those for lesion growth, especially when a is large, although there is a strong correlation between these three variables (Figure 10.8). This suggests that lesion growth occurs after, and presumably as a result of, the expansion in infected phagocyte and bacteria populations.

Overall, these results suggest that high rates of apoptosis and lower rates of phagocyte influx are favourable for keeping the system in the initial plateau phase, during which lesions



(a) Initial plateau level

(b) Breaktimes

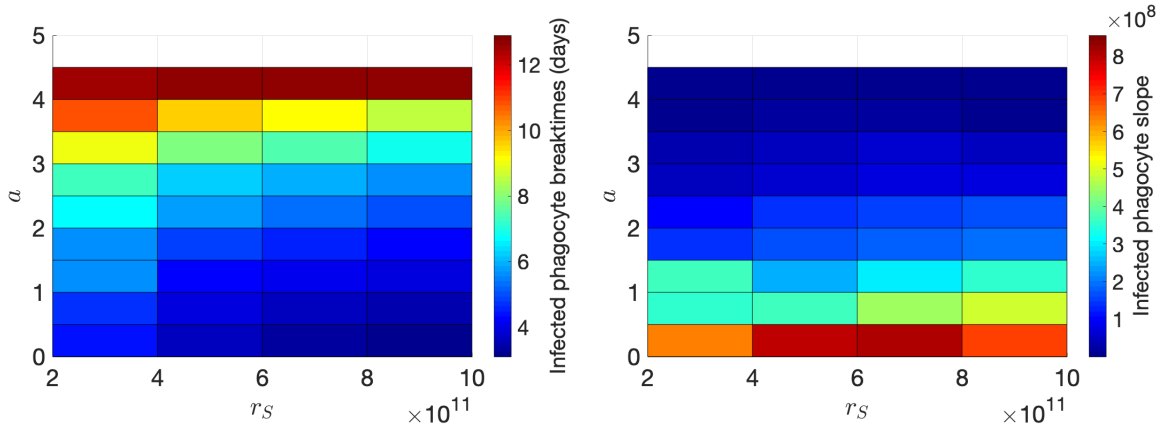


(c) Final slope

(d) Maximum time reached

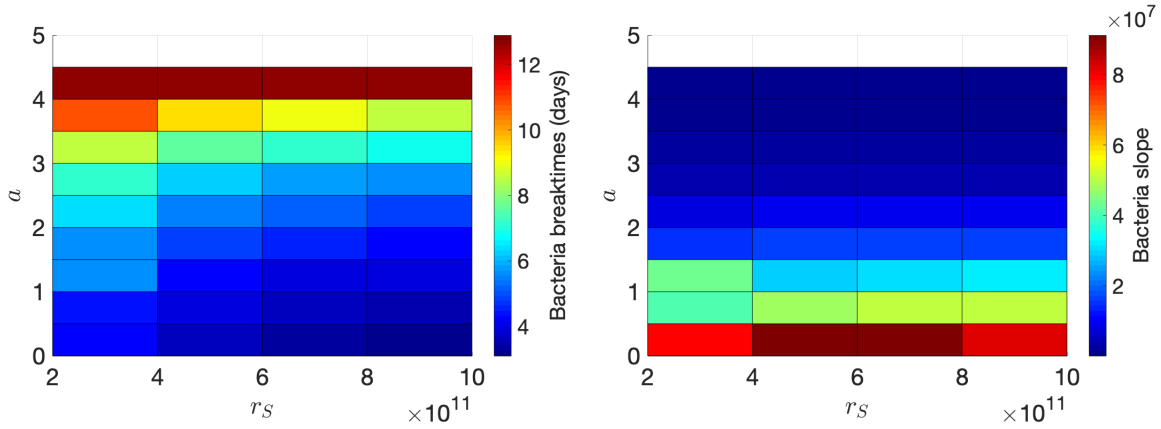
Figure 10.6: Lesion size results while varying apoptosis rate a and rate of phagocyte influx r_S, with no motion of infected phagocytes. Results are extracted from numerical solutions of the PDE model. The initial plateau level is independent of a and r_S. The breaktimes (i.e. the end of the plateau phase) increase with increasing a. The final lesion growth rate increases as a decreases and increases slightly as r_S decreases. The maximum time reached before the solver fails increases as a increases and decreases as r_S increases. This suggests that the lesion remains in the plateau phase for longer for larger a and smaller r_S.

are limited in size, and bacteria and infected phagocyte levels are low; the rates of growth at later times are also slower for higher apoptosis rates. This suggests that higher levels of apoptosis in the system are more likely to lead to stronger control of infection. However, perhaps counter-intuitively, the influx of immune cells appears to be beneficial to the spread of infection, since this supplies more possible hosts for bacteria. Although we do not see a precise analogue of Turing patterning in this system, these results concur with the results from the previous chapter (apoptosis contributes to the stability of the system, phagocyte influx contributes to instability).



(a) Breaktimes (infected phagocytes)

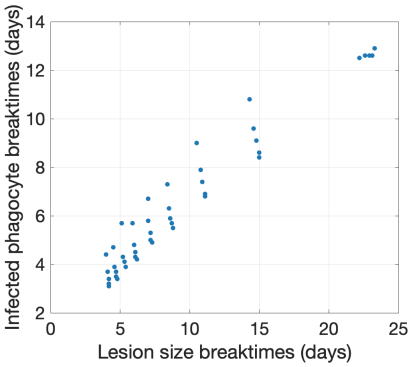
(b) Final slope (infected phagocytes)



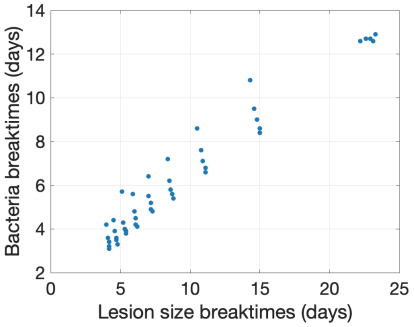
(c) Breaktimes(bacteria)

(d) Final slope (bacteria)

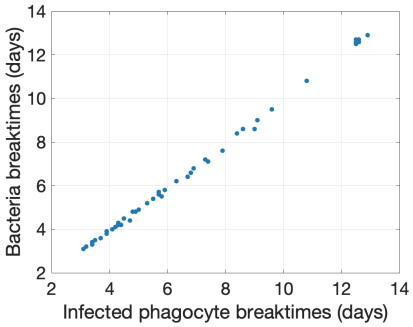
Figure 10.7: Infection level results while varying apoptosis rate a and rate of phagocyte influx r_S, with no motion of infected phagocytes. Results are extracted from numerical solutions of the PDE model. The breaktimes (i.e. the end of the plateau phase) for both infected phagocytes and bacteria increase with increasing a, and decrease slightly with increasing r_S. The final growth rate for both bacteria and phagocytes increases as a decreases. This suggests that the levels of infection remain low for longer, and final growth rates are lower, for larger a.



(a) Lesion size vs infected phagocytes ($\rho = 0.920$)



(b) Lesion size vs bacteria ($\rho = 0.921$)



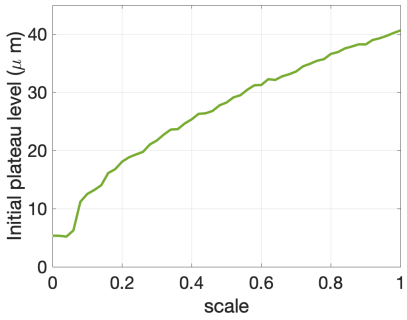
(c) Infected phagocytes vs bacteria ($\rho = 0.998$)

Figure 10.8: Correlations between length of plateau phase for growth in lesion size, total infected phagocyte concentration and total bacteria concentration. Results are extracted from numerical solutions of the PDE model. All three variables are strongly correlated, though not identical (breaktimes for infected phagocyte and bacteria concentrations occur later); Spearman's rho is given for all, $p < 10^{-8}$ in all cases.

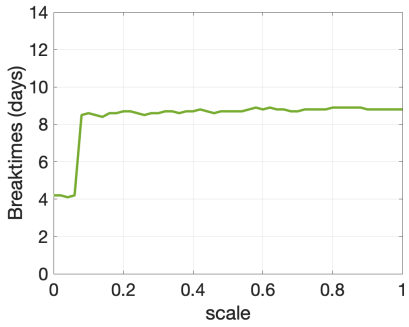
10.3.3 Varying movement parameters for infected phagocytes

Following analysis of the apoptosis and phagocyte influx rates, I then repeated the analysis with different values of a parameter $scale$ relating the movement rates of phagocytes prior to and subsequent to infection; in particular, as in Chapter 9, $D_I = scale \times D_S$ and $\chi_I = scale \times \chi_S$. Unsurprisingly, varying the motion of infected phagocytes has a clear effect on the geometry of lesions. While the relationships established in the previous section between a, r_S and lesion growth still hold, the overall size of the lesion and the rate of growth vary depending on the rate of infected phagocyte motility (example results for $a = 2.5$ and $r_S = 2 \times 10^{11}$ are given in Figure 10.9 but the following conclusions hold for other parametrisations). In general, the size of lesions during the plateau phase increases as the motility of infected phagocytes increases; likewise, with the exception of some simulations run for very low motility, the slope during the growth phase increases as infected phagocyte motility increases. This is unsurprising, since a higher level of motility of infected phagocytes will lead to more spread away from the centre of the lesion area. However, in general, the breaktimes and maximum time reached are approximately constant as the motility of infected phagocytes is varied. Similar results are observed for the total concentrations of bacteria and infected phagocytes; in particular, breaktimes are generally constant as motility of infected phagocytes varies, and rate of growth during the growth phase increases as motility of infected phagocytes increases.

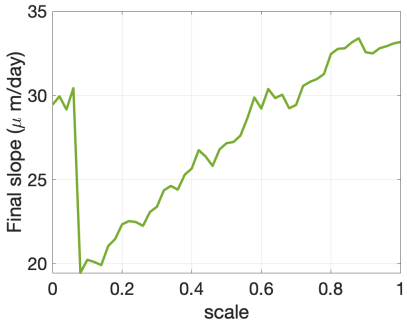
This suggests that decreasing the motility of infected phagocytes has the overall effect of reducing the size of lesions during the plateau phase, and reducing the growth rate of lesions during the growth phase (and likewise, the rate of growth of bacteria and infected phagocyte populations). Note that there are some unusual results for very low infected phagocyte motility (very high final slopes, and relatively low breaktimes and maximum time reached). This may suggest that it is in fact optimal for phagocytes to have low, but non-zero, motility following infection, although it is not clear why this might be the case, and this is possibly an artefact of the simulations. However, it is clear that the reduction of cellular motility observed following infection (see [Man et al., 2014a]) is important for reducing both the spatial spread and overall bacterial burden of the infection.



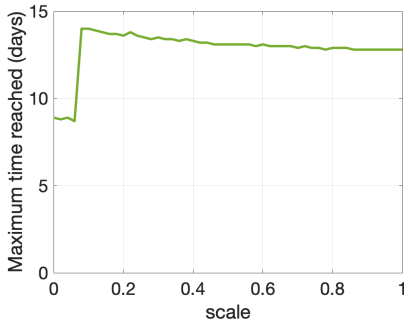
(a) Initial plateau level



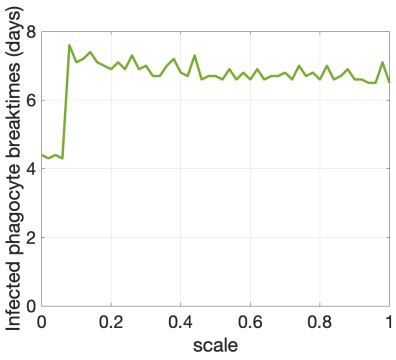
(b) Breaktimes



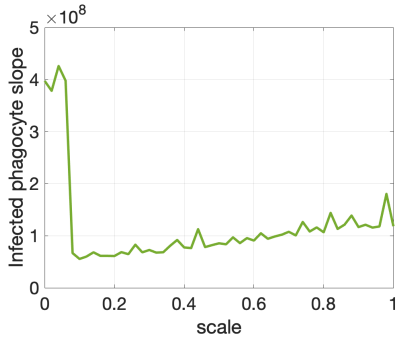
(c) Final slope



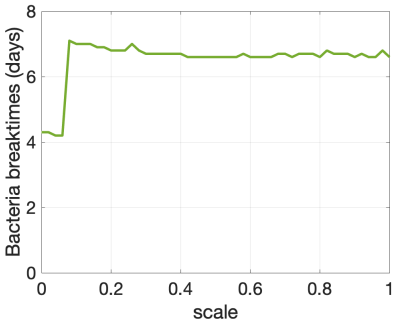
(d) Maximum time reached



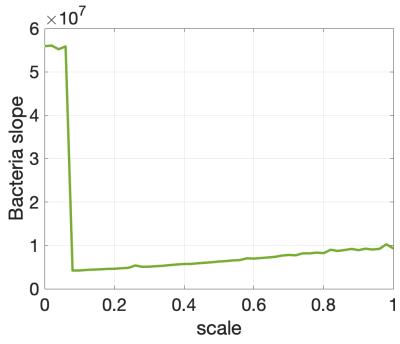
(e) Breaktimes (infected phagocytes)



(f) Final slope (infected phagocytes)



(g) Breaktimes(bacteria)



(h) Final slope (bacteria)

Figure 10.9: Effects of varying motility of infected phagocytes on lesion growth, shown here for apoptosis rate $a = 2.5$ and phagocyte influx rate $r_S = 2 \times 10^{11}$. Results are extracted from numerical solutions of the PDE model. The level of the initial plateau increases as motility scale increases, i.e., lesion size during the plateau phase increases as infected phagocytes become more mobile. Following an initial spike, the final growth rate of lesion area, infected phagocyte population and bacteria population increases as the motility of infected phagocytes increases. In general breaktimes are not affected by the motility of infected cells.

10.4 Discussion

10.4.1 Summary

In this chapter, I have outlined a methodology for simulating the results of the model presented in Chapter 8, and presented the results for a range of different parametrisations.

Even in this simple model, we can see stable lesion-like structures which persist during the first few weeks of infection; this demonstrates that lesions may form simply due to the balance of infection, phagocyte influx, lysis and apoptosis. Infected bacteria and phagocytes are concentrated in round areas which are reminiscent of lesions. However, at later times, these areas grow over time in a linear fashion, with the concentration of bacteria and infected phagocytes in the centre of the lesion increasing exponentially, until the ODE solver used to calculate results fails. This is growth phase is not seen in experimental systems; it is possible that the parameters in the model are such that stability is maintained long enough for another biological process to prevent growth at later times (the arrival of fibroblasts which 'close off' the nascent lesion is a possible explanation). Furthermore, seeding of new lesions does not occur.

I have shown that the length of the initial plateau phase, and the rate of growth of lesions at later times, is considerably slowed when apoptosis is included in the system. On the other hand, increasing the influx rate of phagocytes generally has a destabilising effect, leading to faster onset of lesion growth and more rapid lesion growth and bacterial burden at later times. This suggests that a balance between apoptotic and lytic cell death, as well as influx of uninfected phagocytes, is an important factor in maintaining lesions and reducing overall bacterial load.

I have also shown that increasing motility of infected cells increases the size of lesions early in infection, and the growth rate of lesions, infected phagocytes, and bacteria during the growth phase; this confirms that reduction of motility in infected cells plays an important role in reducing the spatial spread of infection, as well as the bacterial burden, to keep the level and spread of infection at a lower level before other processes intervene.

10.4.2 Further work

The model of lesion formation as it stands may give an account for why stable lesions are observed at early times, up to around 1-2 weeks post-infection, but the sharp rate of growth

of lesions thereafter is unexpected; it is unclear whether other biological systems might be interfering to prevent the growth at later times (for example, the action of fibroblasts, or the initiation of the adaptive immune response), or whether instead this is merely an artefact of the mathematical model which does not arise in the biological system. In order to investigate this further, there are a number of more immediate steps which may be carried out without further alteration of the model itself. Firstly, it may be useful to run the simulations again with a different method of lines scheme, to ensure that the rapid growth at late stages is not an artefact of the solver used here. Furthermore, the standard parametrisation as established in Chapter 8 is mostly based on comparisons to similar systems; in particular, I have focused on varying the apoptosis rate, the rate of phagocyte influx, and the motility of infected phagocytes. A useful next step may be to explore the parameter space of the model more thoroughly, in order to establish whether lesion-like results can be established for any parametrisation. This may also explain why the observed lesions in the solutions for this model are smaller than expected.

It may also be worth investigating different choices of initial condition, especially since at no point did we see new lesion structures forming — in particular, when there are multiple sources of bacteria. This would allow us to determine whether multiple lesions form in this system, and whether they interact.

It is also worth investigating further why the trends in behaviour of the system as motility of infected phagocytes decreases do not hold when motility of infected phagocytes is close to zero. Once again, this may be an artefact arising from inaccuracies in the ODE solving system used, or in breakpoint fitting; however, if these anomalies persist even with different methodologies, this may suggest that it is in fact not optimal for infected phagocytes to become completely immobile.

Finally, so far the results from this modelling project have not been formally compared to experimental data. An ideal dataset would be stained histopathological samples of liver tissue at different timepoints during infection, with lesions identified and their size and distribution through the tissue formally quantified. The availability of a good-quality dataset will be vital to the further development of models in this field.

Improvements to the existing model and suggestions for further work will be discussed in more detail in the following chapter.

Chapter 11

Discussion of tissue-level modelling of lesion formation

In the preceding chapters, I have established a novel spatial PDE model of lesion formation in the liver following *Salmonella* infection; the model focuses particularly on the effects of lytic cell death of infected phagocytes, and the differences in motility between infected and uninfected phagocytes. In this chapter I will explain how the results from this model affect our understanding of lesion formation, and suggest possible extensions to the model which may be useful.

11.1 An account of the formation of lesions during *Salmonella* infections

Both analytical, Turing instability-style analysis and numerical solutions of the spatial PDE model of lesion formation show that round, lesion-like structures can form spontaneously as a result of the interplay between influx of phagocytes, chemotaxis, and a balance between apoptotic and lytic cell death. This is in contrast to the existing models of inflammation, which tend to focus on the overall bacterial burden as a result of infection rather than the spatial structures formed, and in general do not include bacterial release due to phagocyte lysis, or differentiate between the motility of infected and uninfected phagocytes.

A key feature of the simulated solutions to this model is that lesion-like spatial structures are stable in size for the first 1-2 weeks following infection ('plateau phase'). The extracellular bacteria and infected phagocyte population levels are also relatively low during this time.

However, following this initial plateau, there is a rapid (approximately quadratic) increase in infected phagocyte and bacteria concentration, followed by a linear increase in lesion size ('growth phase'). While this growth phase has not been observed experimentally, it tends to occur fairly late in the infection, so it is possible that other biological processes intervene to prevent unlimited growth; however, parametrisations which lead to solutions with longer plateau phases, and lower growth rates during the growth phase, are more likely to reflect overall stability of lesions.

This characterisation of lesion formation is in contrast to previous modelling studies of inflammatory response which do consider spatial structure (in particular, [Pigozzo et al., 2012] and [Bayani et al., 2020a]); these generally show gradual growth of the infected area over time, without a distinct initial lag. It is also notable that the model presented here does not separate the class of phagocytes into neutrophils and active or inactive macrophages (as most existing models do); while the model is therefore unable to capture some of the subtleties of the roles of these different classes of cells, nonetheless the distinction between these roles is not necessary for the development of lesion-like structures.

The length of the initial plateau and the growth rate of the lesion area during the growth phase (as well as the size of infected phagocyte and bacteria populations) are sensitive to the apoptosis rate, the rate of influx of phagocytes, and the motility of infected phagocytes. In particular, the plateau lasts for longer, and the final growth rate is slower, for higher apoptosis rates; this suggests that apoptosis has an important stabilising effect on the system, dampening growth and spatial spread of bacteria. In contrast, the plateau becomes shorter as the rate of phagocyte influx increases. The growth rate during the growth phase is also lower for higher levels of apoptosis. Overall, these results suggest a critical balance between apoptotic and lytic cell death in controlling infection both spatially and in terms of the overall bacterial burden; it also suggests that the influx of phagocytes from the bloodstream must be carefully controlled to avoid providing too many hosts for bacteria.

It is worth considering, especially in the light of the stabilising effect demonstrated for apoptosis, why lytic cell death takes place at all; in this simple model, there is no real benefit for the host for lytic, rather than apoptotic, cell death taking place. It is worth bearing in mind that in biological systems, lytic cell death releases a large concentration of cytokines (and possibly inflammasomes) which promote further inflammation. This was not included in the current model due to lack of quantification of 'burst' release of chemokines, but may be a

useful factor to build into future models.

Another important result is that increasing infected phagocyte motility leads to larger lesions during the plateau phase, and faster growth of lesions (as well as overall infected phagocyte and extracellular bacteria populations) during the growth phase. This supports the hypothesis previously suggested in a study of cellular stiffening carried out by Man *et al.*, that a reduction of motility in phagocytes following infection and inflammasome activation is an important factor in limiting spatial spread and population growth of bacteria and infected cells [Man et al., 2014a]. On the other hand, it is worth noting that the Man *et al.* study shows an increase in the number of lesions in systems with NLRC4 knockout cells (which do not show full inflammasome activation — thus, cell motility is not decreased upon infection of these cells); it is not clear whether the lesions differed in size between wild-type and knock-out samples. This would be worth investigating further, both using mathematical modelling techniques (i.e., developing a model in which seeding of new lesions can take place), as well as carrying out further experimental work to establish explicit measurements of lesion size and distribution for wild-type and NLRC4 knockout systems, at different timepoints during infection.

There are a number of possible reasons for the late time growth phase of lesion size and bacterial burden seen in the model presented here. Firstly, it may simply be an artefact of this particular model. Secondly, other biological process could be intervening before the growth phase can take place; for example, fibroblasts could surround the lesion, preventing further growth. However, this would not prevent the sharp growth of bacteria and infected phagocyte populations. It is also possible that adaptive immune processes are engaged at later time points. Another possible explanation is that the lesion has reached a critical size or concentration, but there is no mechanism for new lesions to be formed instead of bacteria remaining associated with the originating lesion. In particular, note that for the initial conditions used here (one single initial point of bacterial infection) we see the formation of only one lesion. This is in contrast to observed experimental data, in which we see that lesions certainly seed the formation of further spatially distinct lesions. It is possible that the continuous deterministic model presented here is unable to capture instances of bacterial 'escape' from existing lesions, as this may be a somewhat rare event limited to a small number of bacteria. This would relieve the intense concentration of bacteria and infected phagocytes which builds up in the centre of the lesion in the spatial PDE model. This theory

would also explain why the growth phase of the lesion begins after the bacteria and infected phagocyte populations begin to grow. These are certainly considerations which could be investigated in future modelling, either by including more variables in the model (for example, a variable representing fibroblast concentration) or by creating a discrete stochastic version of the model (which may be able to represent bacterial 'escape').

11.2 Further work

There are a number of possible further directions for the work presented here. Many of these centre around possible adjustments which could be made to the model as presented previously; this is a natural stage in the process of model evolution. It will be particularly important to investigate aspects of the results which do not concur with experimental observations — for example, the presence of a growth phase for lesion size and bacteria and infected phagocyte population size.

One key simplification made in the construction of this model was the representation of bacterial dynamics. In particular, I have excluded extracellular bacterial replication in this initial model. In order to render the bacterial dynamics in the model more accurately, the linear extracellular bacterial death rate could be replaced using a logistic growth term, as follows:

$$\frac{\partial b}{\partial t} = D_b \nabla^2 b - \beta S \frac{b}{b + b_{1/2}} + NlI + \underbrace{r_b b \left(1 - \frac{b}{B}\right)}_{\text{Logistic growth}} \tag{11.1}$$

Here, r_b is the initial growth rate and B is the carrying capacity. This representation of the extracellular bacterial population has previously been used elsewhere, and values of r_b and B have been estimated based on experimental data (see, for example, [Schokker et al., 2013]); a term of this form was not included in the initial model since it leads to considerably more complex dynamics than in the simple model presented here. In particular, there would be multiple fixed points in the spatially homogeneous system, and a full bifurcation analysis would be necessary before Turing instability analysis could take place.

It would also be possible to expand the bacterial replication dynamics within infected phagocytes. The class of infected cells could be divided into subclasses $I_1, I_2, I_3, \cdots$ where I_n is the concentration of cells containing exactly n bacteria. The rates of reproduction of bacteria have been shown to be dependent on the current bacterial load in the cell; in par-

ticular, the rate of reproduction when there are n bacteria is $\alpha_n = \alpha_0 e^{-\alpha_e n}$ where α_0 and α_e are constants which have been calculated in the case of *S. enterica* infection [Brown et al., 2006]. It would therefore be fairly straightforward to include a term moving infected cells from the class I_n to I_{n+1} with this rate. This would lead to a much larger model (in particular, with a large number of infected phagocyte classes), and analysis would be correspondingly more laborious; however, this may be useful to gain a deeper understanding of the balance between bacterial replication within the cell, bacteria release upon lysis, and bacterial proliferation outside the cell.

Another possible alteration to the model would be adding another process which stabilises lesion formation at a late stage, to prevent explosive growth occurring. For example, adding fibroblasts to the model would be particularly useful to examine the hypothesis of whether fibroblasts stabilise and surround nascent lesions, and in particular, whether the inclusion of fibroblasts in the model can avert the late growth stage observed in the solutions for the current model.

The input term for phagocytes is currently set to be constant throughout the domain; this is a somewhat unrealistic depiction of the input of phagocytes from the bloodstream, since this is likely to be localised to areas surrounding a blood vessel. The limitation of uninfected phagocyte input to areas where chemokines are present may also be partially responsible for why seeding of new lesions is not observed in this model. The localisation of phagocyte input could be built into a simulation by including a spatial representation of the blood vessel network overlaid on the domain, and allowing phagocytes to enter the domain only from points on this network, possibly at a rate proportional to the concentration of chemokines integrated over the entire domain, or some local neighbourhood of the point of entry (a similar technique has been used elsewhere - see, for example [Pigozzo et al., 2012]). However, such a model would be less amenable to methods of analysis used in Chapter 9.

Another key issue is that the square domain used here, although a simple option which could be used initially, is a poor reflection of the architecture of liver tissues. Rerunning simulations with a more realistic domain architecture may lead to more realistic results - for example, a hexagonal domain would be a better representation of the structure of hepatic lobules. It would also be useful to adjust the simulations to explore the behaviour of the model in three dimensions; once again, a two-dimensional representation was a useful first step but ultimately this is not a good representation of the liver tissue, and adding a third

dimension may substantially change the dynamics of the system and the final spatial pattern formed.

Finally, as described above, some of the processes in the spatial PDE model may be clarified by switching to a discrete stochastic representation. This may be able to capture seeding of new lesions from existing lesions through bacterial 'escape'. Furthermore, the exact size of the lesion may be easier to determine, unlike in the spatial PDE model; for much of the spatial domain the concentrations of all variables is greater than zero for all time, but at many points the concentration is extremely low. This is an artefact of the continuous model, and many of these concentrations are so low that in reality they would correspond to an absence of the corresponding reactant. To delineate the lesion area in the continuous solutions for the PDE model, it was therefore necessary to divide the domain into 'lesion' and 'not lesion' using a somewhat arbitrary threshold.

The discrete stochastic version of the spatial PDE model need not be as complex as the ABMs outlined in Chapter 1 (in particular, the models developed by Shi *et al.* [Shi et al., 2016a, Shi et al., 2016b]). It would be sufficient to track infected and uninfected phagocytes and bacteria as agents in the model, executing a random walk (biased by chemokine levels) on a two-dimensional grid similar to the spatial domain used here and undergoing reactions using a Gillespie-like algorithm similar to that outlined for simulations of the stochastic model of inflammasome formation (see Chapter 5 in particular). Chemokines could be tracked as a background continuous variable, since they are likely to be present at much higher abundances than phagocytes or bacteria. A similar technique to that oulined for the inflammasome formation simulations could be used to understand how varying parameters in the model affects results; in particular, characteristics of each simulation run (for example, number and size of lesions at specific times, or rate of increase of bacteria and infected phagocytes) could be derived for a set of simulations across a range of parametrisations, and the resulting characteristic statistics compared. It would also be possible to prevent overcrowding of cells (which would certainly be an issue with the concentrations arising at late times in the spatial PDE model) by, for example, setting a maximum number of agents allowed to occupy a point in space.

Looking beyond alterations to the spatial PDE model, a further possible extension would be to unite the lesion model presented here with both smaller-scale (within-cell) and larger-scale (organ or host level) modelling. It would also be greatly beneficial to the development

of further models in this field to produce a comprehensive dataset of lesion formation over time; this would allow the model to be parametrised more accurately, as well as providing a concrete reference to test the model's performance. Both of these issues are very broad in scale, and will therefore be considered in more depth in the final discussion chapter.

Chapter 12

Discussion

As outlined in the introduction chapter, the very broadest aim of this research was to improve our understanding of the innate immune response to *Salmonella* infection through the use of mathematical modelling techniques. Within this broad aim, I focused on two specific questions related to distinct, though related, innate immune processes. The first question — how does the inflammasome form? — was addressed in the early chapters of this thesis. The second question — why do lesions form? — was addressed in the latter chapters of this thesis. The results of the research presented go some way towards answering these questions, but they are by no means fully resolved. In the previous discussions (Chapters 7 and 11) I summarised the conclusions from the inflammasome signalling models and the lesion formation model separately, as well as considering the limitations of the approaches used, and suggesting short-term aims for future work. In this final chapter, I will take a broader view; I will discuss the wider implications of the work presented in this thesis as a whole, as well as avenues for future investigation for which this research may serve as a foundation.

12.1 Implications of research

There are clear differences between the broader implications of the two research questions explored in this thesis, mainly due to the fact that the fields of inflammasome signalling and the study of innate immune response to infection at the tissue level are somewhat different.

In particular, in addressing the problem of inflammasome formation, there were virtually no existing models to build on, with the exception of the fairly simplistic model proposed by

Cheng *et al.* [Cheng et al., 2010], whereas numerous models of inflammation due to infection at a tissue level have been constructed already (in particular, the work of Shi *et al.* [Shi et al., 2015a, Shi et al., 2016b, Shi et al., 2016c], which explicitly considers *Salmonella* infection in the liver). Therefore, the construction of a model of inflammasome formation is particularly significant, as it provides an initial framework for the interplay between experimental work in this field and mathematical or computational approaches, while the model of lesion formation builds on existing modelling literature, albeit with a different emphasis to many of the models constructed previously.

The purposes of the inflammasome formation and lesion formation research presented here were also subtly different. The reasons for creating models of inflammasome formation were to integrate existing experimental observations, and to challenge the accepted wisdom which had arisen from experimental work. In contrast, the aim of the lesion formation research was to draw attention to an unusual aspect of this system (namely, the unusual spatial structure of lesions) which had not been fully explained either in experimental accounts or in previous models of this system. The research presented in this thesis has gone some way towards meeting these aims, although the results have not always been entirely conclusive (especially in the case of the inflammasome formation models); nonetheless, modelling of inflammasome formation has highlighted areas of the existing conceptual models of this system which require further experimental investigation, and modelling of *Salmonella* infection at the tissue level has proposed one possible explanation for the spatial structure of lesions (i.e., that these stable structures may form naturally due to the interplay of cell death and phagocyte influx). Both research threads represent the first iteration in a cycle of modelling and experimental work which are mutually beneficial; this cycle will hopefully continue into the future, as well as forming part of a wider movement towards the use of more quantitative approaches in the study of innate immunity.

12.2 Future directions for mathematical modelling in investigating innate immune response to *Salmonella* infection

Although the results presented in this thesis suggest some immediate avenues for further research, which I have outlined in previous chapters, on a broader level there are a number of key gaps in the existing literature on *Salmonella* infection. I will now discuss a number of

areas which would particularly benefit from the use of mathematical modelling and computational techniques.

12.2.1 Mapping regulation of NLR activation

A recurring issue in creating quantitative models of inflammasome formation is that NLR activation and oligomerisation is mediated by numerous other reactants which may control PTMs or bind to inflammasome constituents to prevent further interactions. This is especially an issue in the case of NLRP3 oligomerisation, which is regulated by a particularly complex network of PTMs and chaperones.

Many of the existing studies in this area have focused on a specific element in NLR activation; for example, individual studies have considered the role of NEK7 in stabilising NLRP3 oligomers [He et al., 2016, Schmid-Burgk et al., 2016, Shi et al., 2016b], the role of COPs and POPs in blocking CARD-CARD and PYD-PYD interactions respectively [Stehlik et al., 2003a, Dorfleutner et al., 2007a, Dorfleutner et al., 2007b, Atianand and Harton, 2011], and the role of numerous proteins in phosphorylation and other PTMs [Kelley et al., 2019, Swanson et al., 2019, Yang et al., 2019]. However, there is currently a lack of information regarding how these regulatory processes interact, and together coordinate the oligomerisation of NLRs. This issue was circumvented in the inflammasome models presented in this thesis, by taking pre-established NLR oligomers as inputs for the model, and assuming NLR activation and oligomerisation takes place prior to the processes included in the model. However, this temporal separation of NLR activation and oligomerisation is a major assumption which should certainly be tested further in future models of inflammasome formation.

One possible approach to unifying the existing literature surrounding NLR activation would be to draw on methodology developed in a systems biology context, and integrating results from the experimental literature to build a network model of regulatory processes of NLR activation and oligomerisation. This approach is commonly used in other biological fields (for example, in constructing gene regulatory networks). A network could be developed with nodes corresponding to specific residues in NLRP3 which may be modified by PTMs, individual proteins responsible for coordinating PTMs, stimuli, or PAMPs/DAMPs, and edges corresponding to relationships between these which have been experimentally established (for example, activation or inhibition). A simple binary network may be sufficient initially (with edges being turned ‘on’ or ’off’ corresponding to different cellular conditions, and the re-

sults integrated over the network to give an overall NLRP3 activation level). This network could also form the basis of a large ODE model of interactions between residues in NLRP3, inflammasome stimuli, and proteins responsible for PTMs, although this may require more quantitative data for parametrisation to be possible.

Establishing such a network of control of NLR oligomerisation would be extremely useful for understanding downstream effects on inflammasome formation, but this is also a key part of the innate immune response to infection in its own right; clarifying the relationships between the individual parts of the network which controls NLR oligomerisation may be particularly relevant in understanding the pathology of conditions involving overactivation of NLRs such as CAPS and MAS.

12.2.2 The inflammasome in a cell-wide context

The modelling of inflammasome formation presented here has mainly focused on the overall abundances of inflammasome constituents over time. As for much of the existing literature describing modelling of lesion formation, the focus has been mainly on the outcome of the process based on different model parametrisations, and spatial structure has not been taken into account. In particular, the final position of the inflammasome within the cell and interactions with other organelles has not been considered. However, particularly in the case of the NLRP3 inflammasome, correlations between the position of the inflammasome and other organelles (including Golgi, endoplasmic reticulum and mitochondra) have been noted [Zhou et al., 2011, Misawa et al., 2013, Li et al., 2017], as well as MARK4-driven movement along microtubules [Li et al., 2017]), although results from different studies have been contradictory, and there is no cohesive narrative of subcellular locations of inflammasome constituents throughout the inflammasome formation process. Furthermore, it has been shown that the actin cytoskeleton is reconfigured upon infection of macrophages by *Salmonella* [Méresse et al., 2001, Miao et al., 2003, Poh et al., 2008], and that the actin cytoskeleton is once more reconfigured leading to cellular stiffening as a result of the NLRC4 inflammasome formation process [Man et al., 2014a]. One key question is therefore to what extent interactions with the cytoskeleton (either actin, or microtubules in the case of NLRP3) may contribute towards the formation of the inflammasome, and in particular, whether inflammasome constituents might be actively directed towards a focal point by movement along actin filaments or microtubules.

One way to address this question would be to extend the full stochastic model presented in Chapter 3 to include a spatial element. In particular, the cell could be represented as a lattice, with individual inflammasome constituents (free ASC monomers, NLR oligomers, and clusters) executing random walks on this lattice to simulate diffusion; the same reactions and associated propensities could be used as in the stochastic model from Chapter 3, but reactions would only be permitted to occur when reactants are in close spatial proximity. This simulation could be run many times, possibly with different parametrisations. This simple model with inflammasome formation occurring by diffusion only could be compared to a more complex model, in which interactions with the cytoskeleton are included; this could be as simple as a secondary lattice representing the cytoskeleton network overlaid on the original lattice, which permits more direct movement towards a focal point. Characteristic times similar to those described in this thesis (for example, time of single cluster emergence) could then be derived for simulations of both models, and compared, to determine whether diffusion alone is sufficient for inflammasome formation, or if more directed movement of constituents is necessary.

It would also be possible to include other organelles in these simulations, although this may lead to a considerable increase in complexity; ultimately, it may be useful to use an existing software package developed to simulate within-cell reactions such as VCell [Schaff et al., 1997, Cowan et al., 2012], since these have generally been optimised for efficiency.

12.2.3 A survey of lesion size and distribution in infected tissues

A key limitation in the lesions modelling presented previously is that there is currently a lack of quantitative data describing the formation of lesions in various organs over time, following *Salmonella* infection. Existing datasets tend to consist of average lesion counts over a given area, measured from histological samples taken from livers of infected mice, in which macrophages have been stained; the dataset presented in the previously mentioned study carried out by Man *et al.* is a good example [Man et al., 2014a]. However, in order to compare experimental data directly with the results of lesion modelling, it would be useful for quantifying the size and density of distribution of lesions throughout the tissue. In particular, the consistency of size of lesions over time, and between hosts, would be of interest. It would also be useful to measure the extent to which lesions tend to fall into spatial clusters, or spread out evenly across the tissue, and whether there is any particular colocalisation

with the blood supply.

It would therefore be useful to establish a set of samples taken from mice infected with *Salmonella* at a number of different timepoints throughout infection; as with the Man *et al.* study, it would be especially useful to create a matching dataset for NLRC4 knockout mice, which do not show a decrease in macrophage motility following infection. This would allow us to establish more rigorously the effect of infected cell motility on lesion formation and infection spread.

Naturally, to limit the total number of mice required, judicious experimental design is vital here. The possibility of acquiring data at later times after infection may also be limited, as the health of the mice deteriorates. However, specifically for investigation of plateau and growth phases observed in the spatial PDE lesion formation presented in earlier chapters of this dissertation, it would be particularly useful to take samples early in infection (around 3-5 days), at an intermediate point (7-10 days) and at a later time (approximately 2 weeks) to establish how consistency in lesion size might vary between samples. Samples could be taken from the liver, spleen and bloodstream at each timepoint, rather than just focusing on one organ, which would be extremely useful for establishing results about lesion formation beyond just the liver, and furthermore could be used to parametrise a multi-scale model, as discussed below.

12.2.4 A multi-scale within-host model of *Salmonella* infection

Much of the modelling presented in this thesis has considered the problem of bridging spatial and temporal scales; in particular, creating models of both inflammasome signalling and lesion formation involved integrating results known about individual interactions between reactants to give a full picture of behaviour on a whole-cell or tissue level as appropriate. However, so far I have only considered cell-level and tissue-level responses to infection separately. A natural next step would be to build a model which bridges these two scales, to give a full picture of how inflammasome formation drives lesion formation and infection control on a larger scale. It would even be possible to add an additional larger scale (i.e. interactions between organs) to gain insight into infection clearance on a host-wide level.

There are two possible approaches which could be taken here. The first is to link the spatial PDE model of lesion formation in its continuous form with a model of inflammasome signalling. This could be as simple as introducing time-dependent terms to account for

chemokine release from infected cells, decline in motility of infected cells, or death of infected cells, based on output from the inflammasome model; for example, we could rewrite the spatial PDE model of lesion formation as follows:

$$
\begin{aligned}
\frac{\partial S}{\partial t} &= D_S \nabla^2 S - \nabla \cdot (\chi_S S \nabla c) - \beta S \frac{b}{b + b_{1/2}} + r_S - \gamma_S S \\
\frac{\partial I(\boldsymbol{x}; \tau, t)}{\partial t} &= \begin{cases} D_I(0) \nabla^2 I - \nabla \cdot (\chi_I(0) I \nabla c) + \beta S \frac{b}{b+b_{1/2}} - (l(0) + a(0)) I & \text{if } \tau = 0 \\ D_I(\tau) \nabla^2 I - \nabla \cdot (\chi_I(\tau) I \nabla c) - (l(\tau) + a(\tau)) I & \text{if } \tau > 0 \end{cases} \\
\frac{\partial I(\boldsymbol{x}; \tau, t)}{\partial \tau} &= -\frac{\partial I(\boldsymbol{x}; \tau, t)}{\partial t} \\
\frac{\partial b}{\partial t} &= D_b \nabla^2 b - \beta S \frac{b}{b + b_{1/2}} + \int N(\tau) l(\tau) I d\tau - \gamma_b b \\
\frac{\partial c}{\partial t} &= D_c \nabla^2 c + \int r_c(\tau) I d\tau - \gamma_c c
\end{aligned}
\tag{12.1}
$$

Here, $I = I(\boldsymbol{x}; \tau, t)$ where τ is the time since infection; $\boldsymbol{x}$ and t represent spatial position and time as before. Motility of infected cells (diffusion rate D_I and chemotaxis coefficient χ_I), death rates of infected cells (lysis rate l and apoptosis rate a), rate of chemokine production (r_c) and number of bacteria released during lysis N are all dependent on the time τ since infection. We could determine the functions $D_I(\tau), \chi_I(\tau), l(\tau), a(\tau)$ and $r_c(\tau)$ using a model of inflammasome formation and associated downstream processes; for example, we could extend one of the inflammasome formation models presented in the earlier chapters of this dissertation to include cytokine processing, and use this model to estimate the levels of cytokine production over time to give $r_c(\tau)$; $N(\tau)$ could be estimated building on the work of Brown *et al.* [Brown et al., 2006]. The spatial PDE model of lesion formation as presented in Chapter 8 does not take into account the time taken for different inflammasome-regulated processes to take place, nor does it consider the fact that the levels of (for example) lysis, apoptosis and chemokine production are not constant over time for each cell. We could use this more sophisticated model to gain more insight into how the various downstream effects of inflammasome formation affect infection on the tissue level over time.

An alternative approach would be to construct a cross-scale ABM of infection which could span cell, tissue and whole-host scales. For example, for each cell in a spatial ABM of lesion formation, motility, chemokine release and lysis could be controlled by an individual inflammasome model running within that cell, while other processes (such as diffusion) still take

place on a tissue level only. The total levels of infection resulting from this model could even be used as an input for a larger-scale ABM spanning different organs (for example, liver, spleen and bloodstream, as in the model presented by Grant *et al.* [Grant et al., 2008a]). This would give insight into how processes at the within-cell level control the outcomes of infection at a whole-host level. Furthermore, provided that reliable parameter estimates could be obtained, this would allow us to move from the explanatory modelling approach used throughout this dissertation, with its focus on parsimony and providing broad explanations for experimental observations, to a more specialised model less suitable for broadening scientific understanding, but with more predictive power, which may be particularly useful for simulating infection, inflammation and treatment in biomedical contexts.

The research presented in this thesis therefore has the potential to act as the basis for a wealth of new directions in modelling innate immune response to infection; it is my hope that this will be among many such studies in the years to come.

12.3 Conclusions

As laid out in Chapter 1, this research was guided by two very broad questions about the nature of the innate immune response to infection: why does the inflammasome form? And why do lesions form? This thesis goes some way towards answering these questions. In particular, the first half of the thesis establishes one possible logically coherent account of the early stages of inflammasome formation, and in particular, proposes challenges to some existing popular views of inflammasome formation (for example, by suggesting the need for branch-like reactions between proteins). The models presented also give possible answers for some unanswered questions surrounding the inflammasome formation process (for example, we have seen that variations between inflammasome formation times may be caused by differences of abundance in NLR oligomers).

Meanwhile, the second half of this thesis proposes a mathematical structure within which lesions may arise in tissues simply through the interactions of bacteria, phagocytes and chemokines. Furthermore, the model presented demonstrates how upstream processes such as cellular stiffening, and the balance of apoptotic and lytic cell death, can play important roles in determining the nature of the structure of lesions.

Despite the encouraging progress presented in this thesis, the results presented are only

the first step towards answering our questions about inflammasome formation and lesions. In both cases, the mathematical and computational results presented here will need to be verified through further experimental work. More broadly, however, it is my hope that the research presented here will form a useful foundation for future iterations of the cycle of modelling and experimentation. This interdisciplinary way of working will undoubtedly prove valuable to further deepening our understanding of inflammasome and lesion formation, and more broadly, the innate immune response to infection as a whole.

Bibliography

[Achdou et al., 2013] Achdou, Y., Franchi, B., Marcello, N., and Tesi, M. C. (2013). A qualitative model for aggregation and diffusion of β-amyloid in Alzheimer's disease. *Journal of Mathematical Biology*, 67(6-7):1369–1392.

[Adkar et al., 2017] Adkar, B. V., Manhart, M., Battacharyya, S., Tian, J., Musharbash, M., and Shakhnovich, E. I. (2017). Optimization of lag phase shapes the evolution of a bacterial enzyme. *Nature Ecology and Evolution*, 1(6):149.

[Alam et al., 2015] Alam, M., Deng, X., Philipson, C., Bassaganya-Riera, J., Bisset, K., Carbo, A., Eubank, S., Hontecillas, R., Hoops, S., Mei, Y., Abedi, V., and Marathe, M. (2015). Sensitivity analysis of an ENteric immunity SImulator (ENISI)-based model of immune responses to Helicobacter pylori infection. *PLoS ONE*, 10(9):1–25.

[Aldous, 1999] Aldous, D. J. (1999). Deterministic and stochastic models for coalescence (aggregation and coagulation): A review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48.

[Allen et al., 2009] Allen, I. C., Scull, M. A., Moore, C. B., Holl, E. K., McElvania-TeKippe, E., Taxman, D. J., Guthrie, E. H., Pickles, R. J., and Ting, J. P.-Y. (2009). The NLRP3 Inflammasome Mediates In Vivo Innate Immunity to Influenza A Virus through Recognition of Viral RNA. *Immunity*, 30(4):556–565.

[Alt and Lauffenburger, 1985] Alt, W. and Lauffenburger, D. A. (1985). Transient behaviour of a chemotaxis system modelling certain types of tissue inflammation. *Journal of Mathematical Biology*, 24:691–722.

[Amari and Misra, 1997] Amari, S. V. and Misra, R. B. (1997). Closed-form expressions for distribution of sum of exponential random variables. *IEEE Transactions on Reliability*, 46(4):519–522.

[An, 2001] An, G. (2001). Agent-based computer simulation and SIRS: building a bridge between basic science and clinical trials. *Shock*, 16(4):266–273.

[An and Kulkarni, 2015] An, G. and Kulkarni, S. (2015). An agent-based modeling framework linking inflammation and cancer using evolutionary principles: Description of a generative hierarchy for the hallmarks of cancer and developing a bridge between mechanism and epidemiological data. *Mathematical Biosciences*, 260:16–24.

[Anderson et al., 2008] Anderson, J. M., Rodriguez, A., and Chang, D. T. (2008). Foreign body reaction to biomaterials. *Seminars in Immunology*, 20(2):86–100.

[Andrea et al., 2013] Andrea, S., Horvath, G. L., Monks, B., and Eicke, L. (2013). ASC Speck Formation as a Readout for Inflammasome Activation. In De Nardo Christine M., , and Latz, E., editors, *The Inflammasome: Methods and Protocols*, pages 91–101. Humana Press, Totowa, NJ.

[Arciero et al., 2013] Arciero, J., Ermentrout, G. B., Siggers, R., Afrazi, A., Hackam, D., Vodovotz, Y., and Rubin, J. (2013). Modeling the interactions of bacteria and Toll-like receptor-mediated inflammation in necrotizing enterocolitis. *Journal of Theoretical Biology*, 321:83–99.

[Arciero et al., 2010] Arciero, J. C., Ermentrout, G. B., Upperman, J. S., Vodovotz, Y., and Rubin, J. E. (2010). Using a Mathematical Model to Analyze the Role of Probiotics and Inflammation in Necrotizing Enterocolitis. *PLoS ONE*, 5(4):e10066.

[Arciero et al., 2011] Arciero, J. C., Mi, Q., Branca, M. F., Hackam, D. J., and Swigon, D. (2011). Continuum model of collective cell migration in wound healing and colony expansion. *Biophysical Journal*, 100(3):535–543.

[Atianand and Harton, 2011] Atianand, M. K. and Harton, J. A. (2011). Uncoupling of pyrin-only protein 2 (POP2)-mediated dual regulation of NF-κB and the inflammasome. *Journal of Biological Chemistry*, 286(47):40536–40547.

[Baker et al., 2013] Baker, M., Denman-johnson, S., Brook, B. S., and Owen, M. R. (2013). Mathematical modelling of cytokine-mediated inflammation in rheumatoid arthritis. *Mathematical Medicine and Biology*, 30:311–337.

[Baker et al., 2015] Baker, P. J., Boucher, D., Bierschenk, D., Tebartz, C., Whitney, P. G., D'Silva, D. B., Tanzer, M. C., Monteleone, M., Robertson, A. A., Cooper, M. A., Alvarez-Diaz, S., Herold, M. J., Bedoui, S., Schroder, K., and Masters, S. L. (2015). NLRP3 inflammasome activation downstream of cytoplasmic LPS recognition by both caspase-4 and caspase-5. *European Journal of Immunology*, 45(10):2918–2926.

[Ball et al., 1986] Ball, J. M., Carr, J., and Penrose, O. (1986). The Becker-Döring cluster equations: Basic properties and asymptotic behaviour of solutions. *Communications in Mathematical Physics*, 104(4):657–692.

[Barber et al., 2013] Barber, J., Tronzo, M., Horvat, C. H., Clermont, G., Upperman, J., Vodovotz, Y., and Yotov, I. (2013). A three-dimensional mathematical and computational model of necrotizing enterocolitis. *Journal of Theoretical Biology*, 322:17–32.

[Baringhaus and Franz, 2004] Baringhaus, L. and Franz, C. (2004). On a new multivariate two-sample test. *Journal of Multivariate Analysis*, 88(1):190–206.

[Baroja-Mazo et al., 2014] Baroja-Mazo, A., Martín-Sánchez, F., Gomez, A. I., Martínez, C. M., Amores-Iniesta, J., Compan, V., Barberà-Cremades, M., Yagüe, J., Ruiz-Ortiz, E., Antón, J., Buján, S., Couillin, I., Brough, D., Arostegui, J. I., and Pelegrín, P. (2014). The NLRP3 inflammasome is released as a particulate danger signal that amplifies the inflammatory response. *Nature immunology*, 15(8):738–748.

[Bauernfeind et al., 2009] Bauernfeind, F. G., Horvath, G., Stutz, A., Alnemri, E. S., MacDonald, K., Speert, D., Fernandes-Alnemri, T., Wu, J., Monks, B. G., Fitzgerald, K. A., Hornung, V., and Latz, E. (2009). Cutting Edge: NF-κB Activating Pattern Recognition and Cytokine Receptors License NLRP3 Inflammasome Activation by Regulating NLRP3 Expression. *The Journal of Immunology*, 183(2):787–791.

[Bayani et al., 2020a] Bayani, A., Dunster, J., Crofts, J., and Nelson, M. (2020a). Mechanisms and Points of Control in the Spread of Inflammation : A Mathematical Investigation. *Bulletin of Mathematical Biology*, 82:45.

[Bayani et al., 2020b] Bayani, A., Dunster, J. L., Crofts, J. J., and Nelson, M. R. (2020b). Spatial considerations in the resolution of inflammation : Elucidating leukocyte interactions via an experimentally-calibrated agent-based model. *PLoS Computational Biology*, 16(11):e1008413.

[Beattie et al., 2016] Beattie, L., Sawtell, A., Mann, J., Frame, T. C., Teal, B., de Labastida Rivera, F., Brown, N., Walwyn-Brown, K., Moore, J. W., MacDonald, S., Lim, E. K., Dalton, J. E., Engwerda, C. R., MacDonald, K. P., and Kaye, P. M. (2016). Bone marrow-derived and resident liver macrophages display unique transcriptomic signatures but similar biological functions. *Journal of Hepatology*, 65(4):758–768.

[Becker and Döring, 1935] Becker, R. and Döring, W. (1935). Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. *Annalen der Physik*, 416(8):719–752.

[Bedoya et al., 2007] Bedoya, F., Sandler, L. L., and Harton, J. A. (2007). Pyrin-Only Protein 2 Modulates NF- κB and disrupts ASC:CLR Interactions. *The Journal of Immunology*, 178(6):3837–3845.

[Bertsch et al., 2017] Bertsch, M., Franchi, B., Marcello, N., Tesi, M. C., and Tosin, A. (2017). Alzheimer's disease: A mathematical model for onset and progression. *Mathematical Medicine and Biology*, 34(2):193–214.

[Boucher et al., 2018] Boucher, D., Monteleone, M., Coll, R. C., Chen, K. W., Ross, C. M., Teo, J. L., Gomez, G. A., Holley, C. L., Bierschenk, D., Stacey, K. J., Yap, A. S., Bezbradica, J. S., and Schroder, K. (2018). Caspase-1 self-cleavage is an intrinsic mechanism to terminate inflammasome activity. *Journal of Experimental Medicine*, 215(3):827–840.

[Boxberger et al., 2019] Boxberger, N., Hecker, M., and Zettl, U. K. (2019). Dysregulation of Inflammasome Priming and Activation by MicroRNAs in Human Immune-Mediated Diseases. *The Journal of Immunology*, 202(8):2177–2187.

[Brewer et al., 2019] Brewer, S. M., Brubaker, S. W., and Monack, D. M. (2019). Host inflammasome defense mechanisms and bacterial pathogen evasion strategies. *Current Opinion in Immunology*, 60(Figure 1):63–70.

[Brown et al., 2006] Brown, S. P., Cornell, S. J., Sheppard, M., Grant, A. J., Maskell, D. J., Grenfell, B. T., and Mastroeni, P. (2006). Intracellular demography and the dynamics of Salmonella enterica infections. *PLoS Biology*, 4(11):2091–2098.

[Broz and Dixit, 2016] Broz, P. and Dixit, V. M. (2016). Inflammasomes: mechanism of assembly, regulation and signalling. *Nature Reviews Immunology*, 16(7):407–420.

[Brubaker et al., 2015] Brubaker, S. W., Bonham, K. S., Zanoni, I., and Kagan, J. C. (2015). Innate immune pattern recognition: A cell biological perspective. *Annual Review of Immunology*, 33:257–290.

[Bryan et al., 2010] Bryan, N., Dorfleutner, A., Kramer, S., Yun, C., Rojanasakul, Y., and Stehlik, C. (2010). Differential splicing of the apoptosis-associated speck like protein containing a caspase recruitment domain (ASC) regulates inflammasomes. *Journal of Inflammation*, 7(23).

[Burton, 1977] Burton, J. J. (1977). Nucleation Theory. In Berne, B., editor, *Statistical Mechanics Part A: Equilibrium Techniques*, chapter 6, pages 195–234. Plenum.

[Cai et al., 2014] Cai, X., Chen, J., Xu, H., Liu, S., Jiang, Q.-X., Halfmann, R., and Chen, Z. J. (2014). Prion-like polymerization underlies signal transduction in antiviral immune defense and inflammasome activation. *Cell*, 156(6):1207–1222.

[Calvez et al., 2009] Calvez, V., Lenuzza, N., Oelz, D., Deslys, J. P., Laurent, P., Mouthon, F., and Perthame, B. (2009). Size distribution dependence of prion aggregates infectivity. *Mathematical Biosciences*, 217(1):88–99.

[Canna et al., 2014] Canna, S. W., de Jesus, A. A., Gouni, S., Brooks, S. R., Marrero, B., Liu, Y., DiMattia, M. A., Zaal, K. J. M., Sanchez, G. A. M., Kim, H., Chapelle, D., Plass, N., Huang, Y., Villarino, A. V., Biancotto, A., Fleisher, T. A., Duncan, J. A., O'Shea, J. J., Benseler, S., Grom, A., Deng, Z., Laxer, R. M., and Goldbach-Mansky, R. (2014). An activating NLRC4 inflammasome mutation causes autoinflammation with recurrent macrophage activation syndrome. *Nature genetics*, 46(10):1140–1146.

[Carbo et al., 2013] Carbo, A., Bassaganya-Riera, J., Pedragosa, M., Viladomiu, M., Marathe, M., Eubank, S., Wendelsdorf, K., Bisset, K., Hoops, S., Deng, X., Alam, M., Kronsteiner, B., Mei, Y., and Hontecillas, R. (2013). Predictive Computational Modeling of the Mucosal Immune Responses during Helicobacter pylori Infection. *PLoS ONE*, 8(9).

[Ceresa et al., 2018] Ceresa, M., Olivares, A. L., Noailly, J., and Ballester, M. A. (2018). Coupled immunological and biomechanical model of emphysema progression. *Frontiers in Physiology*, 9(APR):1–16.

[Charu et al., 2017] Charu, V., Zeger, S., Gog, J., Bjørnstad, O. N., Kissler, S., Simonsen, L., Grenfell, B. T., and Viboud, C. (2017). Human mobility and the spatial transmission of influenza in the United States. *PLoS Computational Biology*, 13(2):1–23.

[Cheng et al., 2010] Cheng, J., Waite, A. L., Tkaczyk, E. R., Ke, K., Richards, N., Hunt, A. J., and Gumucio, D. L. (2010). Kinetic properties of ASC protein aggregation in epithelial cells. *Journal of Cellular Physiology*, 222(3):738–747.

[Chu et al., 2015] Chu, L. H., Gangopadhyay, A., Dorfleutner, A., and Stehlik, C. (2015). An updated view on the structure and function of PYRIN domains. *Apoptosis*, 20(2):157–173.

[Cilfone et al., 2013] Cilfone, N. A., Perry, C. R., Kirschner, D. E., and Linderman, J. J. (2013). Multi-Scale Modeling Predicts a Balance of Tumor Necrosis Factor-α and Interleukin-10 Controls the Granuloma Environment during Mycobacterium tuberculosis Infection. *PLoS ONE*, 8(7).

[Cirillo et al., 1998] Cirillo, D. M., Valdivia, R. H., Monack, D. M., and Falkow, S. (1998). Macrophage-dependent induction of the Salmonella pathogenicity island 2 type III secretion system and its role in intracellular survival. *Molecular Microbiology*, 30(1):175–188.

[Clare et al., 2003] Clare, S., Goldin, R., Hale, C., Aspinall, R., Simmons, C., Mastroeni, P., and Dougan, G. (2003). Intracellular adhesion molecule 1 plays a key role in acquired immunity to salmonellosis. *Infection and Immunity*, 71(10):5881–5891.

[Cockrell and An, 2017] Cockrell, C. and An, G. (2017). Sepsis reconsidered: Identifying novel metrics for behavioral landscape characterization with a high-performance computing implementation of an agent-based model. *Journal of Theoretical Biology*, 430:157–168.

[Cohen et al., 2014] Cohen, S. I., Rajah, L., Yoon, C. H., Buell, A. K., White, D. A., Sperling, R. A., Vendruscolo, M., Terentjev, E. M., Dobson, C. M., Weitz, D. A., and Knowles, T. P. (2014). Spatial propagation of protein polymerization. *Physical Review Letters*, 112(9):1–5.

[Conlan, 1997] Conlan, J. W. (1997). Critical roles of neutrophils in host defense against experimental systemic infections of mice by Listeria monocytogenes, Salmonella typhimurium, and Yersinia enterocolitica. *Infection and Immunity*, 65(2):630–635.

[Conlan and North, 1992] Conlan, J. W. and North, R. J. (1992). Early pathogenesis of infection in the liver with the facultative intracellular bacteria Listeria monocytogenes, Francisella tularensis, and Salmonella typhimurium involves lysis of infected hepatocytes by leukocytes. *Infection and Immunity*, 60(12):5164–5171.

[Cooper et al., 2015] Cooper, R. L., Segal, R. A., Diegelmann, R. F., and Reynolds, A. M. (2015). Modeling the effects of systemic mediators on the inflammatory phase of wound healing. *Journal of Theoretical Biology*, 367:86–99.

[Coveney and Wattis, 1996] Coveney, P. V. and Wattis, J. A. D. (1996). Analysis of a Generalized Becker-Doring Model of Self-Reproducing Micelles. *Proceedings: Mathematical, Physical and Engineering Sciences*, 452(1952):2079–2102.

[Cowan et al., 2012] Cowan, A. E., Moraru, I. I., Schaff, J. C., Slepchenko, B. M., and Loew, L. M. (2012). *Spatial Modeling of Cell Signaling Networks*, volume 110. Elsevier Inc.

[Cruz et al., 2007] Cruz, C. M., Rinna, A., Forman, H. J., Ventura, A. L., Persechini, P. M., and Ojcius, D. M. (2007). ATP activates a reactive oxygen species-dependent oxidative stress response and secretion of proinflammatory cytokines in macrophages. *Journal of Biological Chemistry*, 282(5):2871–2879.

[Davies et al., 1999] Davies, S. C., King, J. R., and Wattis, J. A. (1999). Self-similar behaviour in the coagulation equations. *Journal of Engineering Mathematics*, 36(1-2):57–88.

[Davis and Sindi, 2016] Davis, J. K. and Sindi, S. S. (2016). Initial condition of stochastic self-assembly. *Physical Review E*, 93(2):1–5.

[de Alba, 2009] de Alba, E. (2009). Structure and interdomain dynamics of apoptosis-associated speck-like protein containing a CARD (ASC). *Journal of Biological Chemistry*, 284(47):32932–32941.

[de Alba, 2019] de Alba, E. (2019). Structure, interactions and self-assembly of ASC-dependent inflammasomes. *Archives of Biochemistry and Biophysics*, 670(January):15–31.

[de Carvalho et al., 2019] de Carvalho, R. V., Silva, A. L., Santos, L. L., Andrade, W. A., de Sá, K. S., and Zamboni, D. S. (2019). Macrophage priming is dispensable for NLRP3

inflammasome activation and restriction of Leishmania amazonensis replication. *Journal of Leukocyte Biology*, 106:631–640.

[Dick et al., 2016] Dick, M. S., Sborgi, L., Rühl, S., Hiller, S., and Broz, P. (2016). ASC filament formation serves as a signal amplification mechanism for inflammasomes. *Nature communications*, 7:11929.

[Diebolder et al., 2015] Diebolder, C. A., Halff, E. F., Koster, A. J., Huizinga, E. G., and Koning, R. I. (2015). Cryoelectron Tomography of the NAIP5/NLRC4 Inflammasome: Implications for NLR Activation. *Structure*, 23(12):2349–2357.

[Dinarello, 2009] Dinarello, C. A. (2009). Immunological and inflammatory functions of the interleukin-1 family. *Annual Review of Immunology*, 27:519–550.

[Dondolossa et al., 2016] Dondolossa, E., Holzapfel, B. M., Alexander, S., Filippini, S., Hutmacher, D. W., and Friedl, P. (2016). Examination of the foreign body response to biomaterials by nonlinear intravital microscopy. *Nature Biomedical Engineering*, 176(1):139–148.

[Dorfleutner et al., 2007a] Dorfleutner, A., Bryan, N. B., Talbott, S. J., Funya, K. N., Rellick, S. L., Reed, J. C., Shi, X., Rojanasakul, Y., Flynn, D. C., and Stehlik, C. (2007a). Cellular pyrin domain-only protein 2 is a candidate regulator of inflammasome activation. *Infection and Immunity*, 75(3):1484–1492.

[Dorfleutner et al., 2007b] Dorfleutner, A., Talbott, S. J., Bryan, N. B., Funya, K. N., Rellick, S. L., Reed, J. C., Shi, X., Rojanasakul, Y., Flynn, D. C., and Stehlik, C. (2007b). A Shope Fibroma virus PYRIN-only protein modulates the host immune response. *Virus Genes*, 35(3):685–694.

[D'Orsogna et al., 2015] D'Orsogna, M. R., Lei, Q., and Chou, T. (2015). First assembly times and equilibration in stochastic coagulation-fragmentation. *Journal of Chemical Physics*, 143(1).

[Dostert et al., 2008] Dostert, C., Pétrilli, V., Bruggen, R. V., Steele, C., and Brooke, T. (2008). Innate immune activation through Nalp3 inflammasome sensing of asbestos and silica. *Science*, 320(5876):674–677.

[Doumic et al., 2016] Doumic, M., Eugène, S., and Robert, P. (2016). Asymptotics of stochastic protein assembly models. *SIAM Journal on Applied Mathematics*, 76(6):2333–2352.

[Drake, 1972] Drake, R. L. (1972). A general mathematical survey of the coagulation equation. In Hidy, G. M. and Brock, J. R., editors, *Topics in current aerosol research*, pages 204–367. Pergamon Press, New York.

[Druilhe et al., 2001] Druilhe, A., Srinivasula, S. M., Razmara, M., Ahmad, M., and Alnemri, E. S. (2001). Regulation of IL-1beta generation by Pseudo-ICE and ICEBERG, two dominant negative caspase recruitment domain proteins. *Cell death and differentiation*, 8(6):649–657.

[Dunster et al., 2014] Dunster, J., Byrne, H., and King, J. (2014). The Resolution of Inflammation : A Mathematical Model of Neutrophil and Macrophage Interactions. *Bulletin of Mathematical Biology*, 76:1953–1980.

[Dutta-Moscato et al., 2014] Dutta-Moscato, J., Solovyev, A., Mi, Q., Nishikawa, T., and Soto-Gutierrez, A. (2014). A multiscale agent-based in silico model of liver fibrosis progression. *Frontiers in Bioengineering and Biotechnology*, 2:18.

[Engler et al., 2006] Engler, H., Prüss, J., and Webb, G. F. (2006). Analysis of a model for the dynamics of prions II. *Journal of Mathematical Analysis and Applications*, 324(1):98–117.

[Eugène et al., 2016] Eugène, S., Xue, W. F., Robert, P., and Doumic, M. (2016). Insights into the variability of nucleated amyloid polymerization by a minimalistic model of stochastic protein assembly. *Journal of Chemical Physics*, 144(17).

[Evavold et al., 2018] Evavold, C. L., Ruan, J., Tan, Y., Xia, S., Wu, H., and Kagan, J. C. (2018). The Pore-Forming Protein Gasdermin D Regulates Interleukin-1 Secretion from Living Macrophages. *Immunity*, 48(1):35–44.e6.

[Everest et al., 1997] Everest, P., Allen, J., Papakonstantinopoulou, A., Mastroeni, P., Roberts, M., and Dougan, G. (1997). Salmonella typhimurium Infections in Mice Deficient in Interleukin-4 Production: Role of IL-4 in Infection-Associated Pathology. *Journal of Immunology*, 159(4):1820–1827.

[Everest et al., 1998] Everest, P., Roberts, M., and Dougan, G. (1998). Susceptibility to Salmonella typhimurium infection and effectiveness of vaccination in mice deficient in the tumor necrosis factor alpha p55 receptor. *Infection and Immunity*, 66(7):3355–3364.

[Fallahi-Sichani et al., 2010] Fallahi-Sichani, M., Schaller, M. A., Kirschner, D. E., Kunkel, S. L., and Linderman, J. J. (2010). Identification of key processes that control tumor necrosis factor availability in a tuberculosis granuloma. *PLoS Computational Biology*, 6(5):1–19.

[Fernandes-Alnemri et al., 2007] Fernandes-Alnemri, T., Wu, J., Yu, J.-W., Datta, P., Miller, B., Jankowski, W., Rosenberg, S., Zhang, J., and Alnemri, E. (2007). The pyroptosome: a supramolecular assembly of ASC dimers mediating inflammatory cell death via caspase-1 activation. *Cell Death and Differentiation*, 14(9):1590–1604.

[Fernández-Díaz and Gómez-García, 2010] Fernández-Díaz, J. M. and Gómez-García, G. J. (2010). Exact solution of a coagulation equation with a product kernel in the multicomponent case. *Physica D: Nonlinear Phenomena*, 239(5):279–290.

[Fields et al., 1986] Fields, P. I., Swanson, R. V., Haidaris, C. G., and Heffron, F. (1986). Mutants of Salmonella typhimurium that cannot survive within the macrophage are avirulent. *Proceedings of the National Academy of Sciences of the United States of America*, 83(14):5189–5193.

[Fornari et al., 2019] Fornari, S., Schäfer, A., Jucker, M., Goriely, A., and Kuhl, E. (2019). Prion-like spreading of Alzheimer's disease within the brain's connectome. *Journal of the Royal Society Interface*, 16(159).

[Fornari et al., 2020] Fornari, S., Schäfer, A., Kuhl, E., and Goriely, A. (2020). Spatially-extended nucleation-aggregation-fragmentation models for the dynamics of prion-like neurodegenerative protein-spreading in the brain and its connectome. *Journal of Theoretical Biology*, 486:110102.

[Fortin, 2019] Fortin, J. Y. (2019). Critical properties of cluster size distribution in an asymmetric diffusion-aggregation model. *Physical Review E*, 100(5):1–8.

[Franchi et al., 2009] Franchi, L., Eigenbrod, T., and Núñez, G. (2009). Cutting Edge: TNF-α Mediates Sensitization to ATP and Silica via the NLRP3 Inflammasome in the Absence of Microbial Stimulation. *The Journal of Immunology*, 183(2):792–796.

[Franklin et al., 2014] Franklin, B. S., Bossaller, L., Nardo, D. D., Ratter, J. M., Stutz, A., Engels, G., Brenker, C., Nordhoff, M., Sandra, R., Al-amoudi, A., Mangan, M., Zimmer, S., Monks, B., Fricke, M., Schmidt, R. E., Espevik, T., Jones, B., Andrew, G., Hansbro, P. M., Busto, P., Marshak-rothstein, A., Aguzzi, A., Kastenmüller, W., and Latz, E. (2014). ASC has extracellular and prionoid activities that propagate inflammation. *Nature Immunology*, 15(8):727–737.

[Franklin et al., 2018] Franklin, B. S., Latz, E., and Schmidt, F. I. (2018). The intra- and extracellular functions of ASC specks. *Immunological Reviews*, 281:74–87.

[Franz, 2019] Franz, C. (2019). cramer: Multivariate Nonparametric Cramer-Test for the Two-Sample-Problem.

[Frevert et al., 2006] Frevert, U., Usynin, I., Baer, K., and Klotz, C. (2006). Nomadic or sessile: Can Kupffer cells function as portals for malaria sporozoites to the liver? *Cellular Microbiology*, 8(10):1537–1546.

[Fronczak et al., 2019] Fronczak, A., Łepek, M., Kukliński, P., and Fronczak, P. (2019). Coagulation with product kernel and arbitrary initial conditions: Exact kinetics within the Marcus-Lushnikov framework. *Physical Review E*, 99(1):1–10.

[Furter et al., 2019] Furter, M., Sellin, M. E., Hansson, G. C., and Hardt, W. D. (2019). Mucus Architecture and Near-Surface Swimming Affect Distinct Salmonella Typhimurium Infection Patterns along the Murine Intestinal Tract. *Cell Reports*, 27(9):2665–2678.e3.

[Gaidt et al., 2016] Gaidt, M. M., Ebert, T. S., Chauhan, D., Schmidt, T., Schmid-Burgk, J. L., Rapino, F., Robertson, A. A. B., Cooper, M. A., Graf, T., and Hornung, V. (2016). Human Monocytes Engage an Alternative Inflammasome Pathway. *Immunity*, 44(4):833–846.

[Galina et al., 2002] Galina, H., Lechowicz, J. B., and Walczak, M. (2002). Kinetic modeling of hyperbranched polymerization involving an AB2 monomer reacting with substitution effect. *Macromolecules*, 35(8):3253–3260.

[Gerisch and Chaplain, 2006] Gerisch, A. and Chaplain, M. A. (2006). Robust numerical methods for taxis-diffusion-reaction systems: Applications to biomedical problems. *Mathematical and Computer Modelling*, 43(1-2):49–75.

[Gerisch et al., 2001] Gerisch, A., Griffiths, D. F., Weiner, R., and Chaplain, M. A. (2001). A positive splitting method for mixed hyperbolic-parabolic systems. *Numerical Methods for Partial Differential Equations*, 17(2):152–168.

[Gerisch and Verwer, 2002] Gerisch, A. and Verwer, J. G. (2002). Operator splitting and approximate factorization for taxis-diffusion-reaction models. *Applied Numerical Mathematics*, 42(1-3):159–176.

[Gillespie, 1972] Gillespie, D. T. (1972). The stochastic coalescence model for cloud droplet growth. *Journal of the Atmospheric Sciences*, 29:1496–1510.

[Gillespie, 1976] Gillespie, D. T. (1976). A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. *Journal of Computational Physics*, 22:403–434.

[Gog et al., 2012] Gog, J. R., Murcia, A., Osterman, N., Restif, O., McKinley, T. J., Sheppard, M., Achouri, S., Wei, B., Mastroeni, P., Wood, J. L. N., Maskell, D. J., Cicuta, P., and Bryant, C. E. (2012). Dynamics of Salmonella infection of macrophages at the single cell level. *Journal of The Royal Society Interface*, 9(75):2696–2707.

[Gong et al., 2021] Gong, Q., Robinson, K., Xu, C., Huynh, P. T., Chong, K. H. C., Tan, E. Y. J., Zhang, J., Boo, Z. Z., Teo, D. E. T., Lay, K., Zhang, Y., Lim, J. S. Y., Goh, W. I., Wright, G., Zhong, F. L., Reversade, B., and Wu, B. (2021). Structural basis for distinct inflammasome complex assembly by human NLRP1 and CARD8. *Nature Communications*, 12(1).

[Grant et al., 2008a] Grant, A. J., Restif, O., McKinley, T. J., Sheppard, M., Maskell, D. J., and Mastroeni, P. (2008a). Modelling within-host spatiotemporal dynamics of invasive bacterial disease. *PLoS Biology*, 6(4):757–770.

[Grant et al., 2008b] Grant, A. J., Sheppard, M., Deardon, R., Brown, S. P., Foster, G., Bryant, C. E., Maskell, D. J., and Mastroeni, P. (2008b). Caspase-3-dependent phagocyte death during systemic Salmonella enterica serovar Typhimurium infection of mice. *Immunology*, 125(1):28–37.

[Greer et al., 2006] Greer, M. L., Pujo-Menjouet, L., and Webb, G. F. (2006). A mathematical analysis of the dynamics of prion proliferation. *Journal of Theoretical Biology*, 242(3):598–606.

[Groß et al., 2016] Groß, C. J., Mishra, R., Schneider, K. S., Médard, G., Wettmarshausen, J., Dittlein, D. C., Shi, H., Gorka, O., Koenig, P. A., Fromm, S., Magnani, G., Ćiković, T., Hartjes, L., Smollich, J., Robertson, A. A., Cooper, M. A., Schmidt-Supprian, M., Schuster, M., Schroder, K., Broz, P., Traidl-Hoffmann, C., Beutler, B., Kuster, B., Ruland, J., Schneider, S., Perocchi, F., and Groß, O. (2016). K+ Efflux-Independent NLRP3 Inflammasome Activation by Small Molecules Targeting Mitochondria. *Immunity*, 45(4):761–773.

[Gross et al., 2009] Gross, O., Poeck, H., Bscheider, M., Dostert, C., Hannesschläger, N., Endres, S., Hartmann, G., Tardivel, A., Schweighoffer, E., Tybulewicz, V., Mocsai, A., Tschopp, J., and Ruland, J. (2009). Syk kinase signalling couples to the Nlrp3 inflammasome for anti-fungal host defence. *Nature*, 459(7245):433–436.

[Gueron and Levin, 1995] Gueron, S. and Levin, S. A. (1995). The dynamics of group formation. *Mathematical Biosciences*, 128(1-2):243–264.

[Hafner-Bratkovič et al., 2018] Hafner-Bratkovič, I., Sušjan, P., Lainšček, D., Tapia-Abellán, A., Cerović, K., Kadunc, L., Angosto-Bazarra, D., Pelegrín, P., and Jerala, R. (2018). NLRP3 lacking the leucine-rich repeat domain can be fully activated via the canonical inflammasome pathway. *Nature Communications*, 9(1).

[Halff et al., 2012] Halff, E. F., Diebolder, C. A., Versteeg, M., Schouten, A., Brondijk, T. H. C., and Huizinga, E. G. (2012). Formation and structure of a NAIP5-NLRC4 inflammasome induced by direct interactions with conserved N- and C-terminal regions of flagellin. *Journal of Biological Chemistry*, 287(46):38460–38472.

[Han et al., 2019] Han, X., Sun, S., Sun, Y., Song, Q., Zhu, J., Song, N., Chen, M., Sun, T., Xia, M., Ding, J., Lu, M., Yao, H., and Hu, G. (2019). Small molecule-driven NLRP3 inflammation inhibition via interplay between ubiquitination and autophagy: implications for Parkinson disease. *Autophagy*, 15(11):1860–1881.

[Hao et al., 2014] Hao, W., Crouser, E. D., and Friedman, A. (2014). Mathematical model of sarcoidosis. *Proceedings of the National Academy of Sciences of the United States of America*, 111(45):16065–16070.

[Hao and Friedman, 2014] Hao, W. and Friedman, A. (2014). The LDL-HDL profile determines the risk of atherosclerosis: A mathematical model. *PLoS ONE*, 9(3).

[He et al., 2015] He, W. T., Wan, H., Hu, L., Chen, P., Wang, X., Huang, Z., Yang, Z. H., Zhong, C. Q., and Han, J. (2015). Gasdermin D is an executor of pyroptosis and required for interleukin-1β secretion. *Cell Research*, 25(12):1285–1298.

[He et al., 2013] He, Y., Franchi, L., and Núñez, G. (2013). TLR Agonists Stimulate Nlrp3-Dependent IL-1β Production Independently of the Purinergic P2X7 Receptor in Dendritic Cells and In Vivo. *The Journal of Immunology*, 190(1):334–339.

[He et al., 2016] He, Y., Zeng, M. Y., Yang, D., Motro, B., and Núñez, G. (2016). NEK7 is an essential mediator of NLRP3 activation downstream of potassium efflux. *Nature*, 530(7590):354–357.

[Helaine et al., 2010] Helaine, S., Thompson, J. A., Watson, K. G., Liu, M., Boyle, C., and Holden, D. W. (2010). Dynamics of intracellular bacterial replication at the single cell level. *Proceedings of the National Academy of Sciences of the United States of America*, 107(8):3746–3751.

[Helal et al., 2013] Helal, M., Hingant, E., Pujo-Menjouet, L., and Webb, G. F. (2013). Alzheimer's disease: analysis of a mathematical model incorporating the role of prions. *Journal of Mathematical Biology*, 69(5):1207–1235.

[Hendriks and Ernst, 1984] Hendriks, E. M. and Ernst, M. H. (1984). Exactly soluble addition and condensation models in coagulation kinetics. *Journal of Colloid And Interface Science*, 97(1):176–194.

[Herrero et al., 2000] Herrero, M. A., Velázquez, J. J. L., and Wrzosek, D. (2000). Sol – gel transition in a coagulation – diffusion model. *Physica D*, 141:221–247.

[Higham, 2021] Higham, N. (2021). Matrix exponential times a vector.

[Hirose et al., 1999] Hirose, K., Nishimura, H., Matsuguchi, T., and Yoshikai, Y. (1999). Endogenous IL-15 might be responsible for early protection by natural killer cells against infection with an avirulent strain of Salmonella choleraesuis in mice. *Journal of Leukocyte Biology*, 66(3):382–390.

[Holcman, 2017] Holcman, D., editor (2017). *Stochastic processes, multiscale modeling, and numerical methods for computational cellular biology*. Springer.

[Hollingsworth et al., 2021] Hollingsworth, L. R., David, L., Li, Y., Griswold, A. R., Ruan, J., Sharif, H., Fontana, P., Orth-He, E. L., Fu, T. M., Bachovchin, D. A., and Wu, H. (2021). Mechanism of filament formation in UPA-promoted CARD8 and NLRP1 inflammasomes. *Nature Communications*, 12(1):1–13.

[Hopkins et al., 2000] Hopkins, S. A., Niedergang, F., Corthesy-Theulaz, I. E., and Kraehenbuhl, J. P. (2000). A recombinant Salmonella typhimurium vaccine strain is taken up and survives within murine Peyer's patch dendritic cells. *Cellular Microbiology*, 2(1):59–68.

[Hormaeche and Mastroeni, 1990] Hormaeche, C. E. and Mastroeni, P. (1990). T cells do not mediate the initial suppression of a salmonella infection in the RES. *Immunology*, 70:247–250.

[Hornung et al., 2008] Hornung, V., Bauernfeind, F., Halle, A., Samstad, E. O., Kono, H., Rock, K. L., Fitzgerald, K. A., and Latz, E. (2008). Silica crystals and aluminum salts activate the NALP3 inflammasome through phagosomal destabilization. *Nature Immunology*, 9(8):847–856.

[Hoss et al., 2019] Hoss, F., Mueller, J. L., Rojas Ringeling, F., Rodriguez-Alcazar, J. F., Brinkschulte, R., Seifert, G., Stahl, R., Broderick, L., Putnam, C. D., Kolodner, R. D., Canzar, S., Geyer, M., Hoffman, H. M., and Latz, E. (2019). Alternative splicing regulates stochastic NLRP3 activity. *Nature Communications*, 10(1).

[Hozé and Holcman, 2014] Hozé, N. and Holcman, D. (2014). Modeling capsid kinetics assembly from the steady state distribution of multi-sizes aggregates. *Physics Letters, Section A: General, Atomic and Solid State Physics*, 378(5-6):531–534.

[Hu and Othmer, 2011] Hu, J. and Othmer, H. G. (2011). A theoretical analysis of filament length fluctuations in actin and other polymers. *Journal of Mathematical Biology*, 63(6):1001–1049.

[Hu et al., 2013] Hu, Z., Yan, C., Liu, P., Huang, Z., Ma, R., Zhang, C., Wang, R., Zhang, Y., Martinon, F., Miao, D., Deng, H., Wang, J., Chang, J., and Chai, J. (2013). Crystal Structure of NLRC4 Reveals Its Autoinhibition Mechanism. *Science*, 341(6142):172–175.

[Hu et al., 2015] Hu, Z., Zhou, Q., Zhang, C., Fan, S., Cheng, W., Zhao, Y., Shao, F., Wang, H.-w., Sui, S.-f., and Chai, J. (2015). Structural and biochemical basis for induced self-propagation of NLRC4. *Science*, 350(6259):399–404.

[Humke et al., 2000] Humke, E. W., Shriver, S. K., Starovasnik, M. A., Fairbrother, W. J., and Dixit, V. M. (2000). ICEBERG : A Novel Inhibitor of Interleukin-1ß Generation. *Cell*, 103:99–111.

[Hussain et al., 2014] Hussain, M., Kumar, J., Peglow, M., and Tsotsas, E. (2014). On two-compartment population balance modeling of spray fluidized bed agglomeration. *Computers and Chemical Engineering*, 61:185–202.

[Ibraguimov et al., 2012] Ibraguimov, A., Owens, L., Su, J., and Tang, L. (2012). Stability Analysis of a Model for Foreign Body Fibrotic Reactions. *Computational and Mathematical Methods in Medicine*, 2012:809864.

[Jeldres et al., 2018] Jeldres, R. I., Fawell, P. D., and Florio, B. J. (2018). Population balance modelling to describe the particle aggregation process: A review. *Powder Technology*, 326:190–207.

[Jullien, 1987] Jullien, R. (1987). Aggregation phenomena and fractal aggregates. *Contemporary Physics*, 28(5):477–493.

[Junatas et al., 2017] Junatas, K. L., Tonar, Z., Kubíková, T., Liška, V., Pálek, R., Mik, P., Králíčková, M., and Witter, K. (2017). Stereological analysis of size and density of hepatocytes in the porcine liver. *Journal of Anatomy*, 230(4):575–588.

[Kagaya et al., 1989] Kagaya, K., Watanabe, K., and Fukazawa, Y. (1989). Capacity of recombinant gamma interferon to activate macrophages for Salmonella-killing activity. *Infection and Immunity*, 57(2):609–615.

[Kalapothakis et al., 2015] Kalapothakis, J. M., Morris, R. J., Szavits-Nossan, J., Eden, K., Covill, S., Tabor, S., Gillam, J., Barran, P. E., Allen, R. J., and MacPhee, C. E. (2015). A kinetic study of ovalbumin fibril formation: The importance of fragmentation and end-joining. *Biophysical Journal*, 108(9):2300–2311.

[Kanneganti et al., 2006] Kanneganti, T.-D., Body-Malapel, M., Amer, A., Park, J.-H., Whitfield, J., Franchi, L., Taraporewala, Z. F., Miller, D., Patton, J. T., Inohara, N., and Núñez, G.

(2006). Critical Role for Cryopyrin/Nalp3 in Activation of Caspase-1 in Response to Viral Infection and Double-stranded RNA. *Journal of Biological Chemistry*, 281(48):36560–36568.

[Kashchiev, 2018] Kashchiev, D. (2018). Growth probability and formation time of the individual Oosawa-Kasai protein fibril. *Physical Review E*, 98(1):8–15.

[Kayagaki et al., 2011] Kayagaki, N., Warming, S., Lamkanfi, M., Walle, L. V., Louie, S., Dong, J., Newton, K., Qu, Y., Liu, J., Heldens, S., Zhang, J., Lee, W. P., Roose-Girma, M., and Dixit, V. M. (2011). Non-canonical inflammasome activation targets caspase-11. *Nature*, 479(7371):117–121.

[Ke et al., 2003] Ke, J. H., Lin, Z. Q., and Wang, X. H. (2003). Exact solution of the cluster size distribution for multi-polymer coagulation process. *Chinese Physics Letters*, 20(1):151–154.

[Keller and Segel, 1970] Keller, E. F. and Segel, L. A. (1970). Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26(3):399–415.

[Keller and Segel, 1971] Keller, E. F. and Segel, L. A. (1971). Model for Chemotaxis. *Journal of Theoretical Biology*, 30:225–234.

[Kelley et al., 2019] Kelley, N., Jeltema, D., Duan, Y., and He, Y. (2019). The NLRP3 Inflammasome: An Overview of Mechanisms of Activation and Regulation. *International Journal of Molecular Sciences*, 20:3328.

[Kim et al., 2009] Kim, Y., Lawler, S., Nowicki, M. O., Chiocca, E. A., and Friedman, A. (2009). A mathematical model for pattern formation of glioma cells outside the tumor spheroid core. *Journal of Theoretical Biology*, 260(3):359–371.

[Kim et al., 2011] Kim, Y., Roh, S., Lawler, S., and Friedman, A. (2011). miR451 and AMPK mutual antagonism in glioma cell migration and proliferation: A mathematical model. *PLoS ONE*, 6(12).

[Krapivsky, 2015] Krapivsky, P. L. (2015). Aggregation driven by a localized source. *Journal of Physics A: Mathematical and Theoretical*, 48(24).

[Kumar et al., 2004] Kumar, R., Clermont, G., Vodovotz, Y., and Chow, C. C. (2004). The dynamics of acute inflammation. *Journal of Theoretical Biology*, 230:145–155.

[Kuri et al., 2017] Kuri, P., Schieber, N. L., Thumberger, T., Wittbrodt, J., Schwab, Y., and Leptin, M. (2017). Dynamics of in vivo ASC speck formation. *The Journal of Cell Biology*, 216(9):2891.

[Lamkanfi et al., 2004] Lamkanfi, M., Denecker, G., Kalai, M., D'Hondt, K., Meeus, A., Declercq, W., Saelens, X., and Vandenabeele, P. (2004). INCA, a novel human caspase recruitment domain protein that inhibits interleukin-1b generation. *Journal of Biological Chemistry*, 279(50):51729–51738.

[Lauffenburger and Keller, 1979] Lauffenburger, D. and Keller, K. H. (1979). Effects of leukocyte random motility and chemotaxis in tissue inflammatory response. *Journal of Theoretical Biology*, 81(3):475–503.

[Lauffenburger and Kennedy, 1981] Lauffenburger, D. A. and Kennedy, C. R. (1981). Analysis of a lumped model for tissue inflammation dynamics. *Mathematical Biosciences*, 53:189–221.

[Lauffenburger and Kennedy, 1983] Lauffenburger, D. A. and Kennedy, C. R. (1983). Localized bacterial infection in a distributed model for tissue inflammation. *Journal of Mathematical Biology*, 16:141–163.

[Laurenzi et al., 2002] Laurenzi, I. J., Bartels, J. D., and Diamond, S. L. (2002). A general algorithm for exact simulation of multicomponent aggregation processes. *Journal of Computational Physics*, 177(2):418–449.

[Lazzari et al., 2014] Lazzari, S., Hamzehlou, S., Reyes, Y., Leiza, J. R., Costa, M. R. P., Dias, R. C., and Storti, G. (2014). Bulk Crosslinking Copolymerization: Comparison of Different Modeling Approaches. *Macromolecular Reaction Engineering*, 8(10):678–695.

[Lechtenberg et al., 2014] Lechtenberg, B. C., Mace, P., and Riedl, S. (2014). Structural mechanisms in NLR inflammasome signaling. *Current Opinion in Structural Biology*, 29:17–25.

[Lee et al., 2019] Lee, C., Do, H. T. T., Her, J., Kim, Y., Seo, D., and Rhee, I. (2019). Inflammasome as a promising therapeutic target for cancer. *Life Sciences*, 231(May).

[Lee, 2000] Lee, M. H. (2000). On the Validity of the Coagulation Equation and the Nature of Runaway Growth. *Icarus*, 143(1):74–86.

[Lee et al., 2001] Lee, S. H., Stehlik, C., and Reed, J. C. (2001). COP, a Caspase Recruitment Domain-containing Protein and Inhibitor of Caspase-1 Activation Processing. *Journal of Biological Chemistry*, 276(37):34495–34500.

[Li et al., 2008] Li, N. Y., Verdolini, K., Clermont, G., Mi, Q., Rubinstein, E. N., Hebda, P. A., and Vodovotz, Y. (2008). A patient-specific in silico model of inflammation and healing tested in acute vocal fold injury. *PLoS ONE*, 3(7).

[Li et al., 2017] Li, X., Thome, S., Ma, X., Amrute-Nayak, M., Finigan, A., Kitt, L., Masters, L., James, J. R., Shi, Y., Meng, G., and Mallat, Z. (2017). MARK4 regulates NLRP3 positioning and inflammasome activation through a microtubule-dependent mechanism. *Nature Communications*, 8(May):1–13.

[Li et al., 2010] Li, X.-D., Lin, Z.-Q., Mei-Xia, S., and Ke, J.-H. (2010). Competing role of catalysis-coagulation and catalysis-fragmentation in kinetic aggregation behaviours. *Chinese Physics B*, 19(12):128201.

[Li et al., 2018] Li, Y., Fu, T. M., Lu, A., Witt, K., Ruan, J., Shen, C., and Wu, H. (2018). Cryo-EM structures of ASC and NLRC4 CARD filaments reveal a unified mechanism of nucleation and activation of caspase-1. *Proceedings of the National Academy of Sciences of the United States of America*, 115(43):10845–10852.

[Li and Roberts, 2009] Li, Y. and Roberts, C. J. (2009). Lumry-eyring nucleated-polymerization model of protein aggregation kinetics. 2. Competing growth via condensation and chain polymerization. *Journal of Physical Chemistry B*, 113(19):7020–7032.

[Lightfield et al., 2011] Lightfield, K. L., Persson, J., Trinidad, N. J., Brubaker, S. W., Kofoed, E. M., Sauer, J. D., Dunipace, E. A., Warren, S. E., Miao, E. A., and Vance, R. E. (2011). Differential requirements for NAIP5 in activation of the NLRC4 inflammasome. *Infection and Immunity*, 79(4):1606–1614.

[Lindgren et al., 1996] Lindgren, S. W., Stojiljkovic, I., and Heffron, F. (1996). Macrophage killing is an essential virulence mechanism of Salmonella typhimurium. *Proceedings of the National Academy of Sciences of the United States of America*, 93(9):4197–4201.

[Liu et al., 2014] Liu, T., Yamaguchi, Y., Shirasaki, Y., Shikada, K., Yamagishi, M., Hoshino, K., Kaisho, T., Takemoto, K., Suzuki, T., Kuranaga, E., Ohara, O., and Miura, M. (2014).

Single-cell imaging of caspase-1 dynamics reveals an all-or-none inflammasome signaling response. *Cell Reports*, 8(4):974–982.

[Liu et al., 2017] Liu, W., Liu, X., Li, Y., Zhao, J., Liu, Z., Hu, Z., Wang, Y., Yao, Y., Miller, A. W., Su, B., Cookson, M. R., Li, X., and Kang, Z. (2017). LRRK2 promotes the activation of NLRC4 inflammasome during Salmonella Typhimurium infection. *The Journal of Experimental Medicine*, 214(10):3051–3066.

[Liu et al., 2016] Liu, X., Zhang, Z., Ruan, J., Pan, Y., Magupalli, V. G., Wu, H., and Lieberman, J. (2016). Inflammasome-activated gasdermin D causes pyroptosis by forming membrane pores. *Nature*, 535(7610).

[Lombardo et al., 2017] Lombardo, M. C., Barresi, R., Bilotta, E., Gargano, F., Pantano, P., and Sammartino, M. (2017). Mathematical Biology Demyelination patterns in a mathematical model. *Journal of Mathematical Biology*, 75(2):373–417.

[López-López et al., 2005] López-López, J. M., Moncho-Jordá, A., Schmitt, A., and Hidalgo-Álvarez, R. (2005). Formation and structure of stable aggregates in binary diffusion-limited cluster-cluster aggregation processes. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 72(3):1–9.

[Lu et al., 2014] Lu, A., Magupalli, V. G., Ruan, J., Yin, Q., Atianand, M. K., Vos, M. R., Schröder, G. F., Fitzgerald, K. A., Wu, H., and Egelman, E. H. (2014). Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes. *Cell*, 156(6):1193–1206.

[Lushnikov, 1978] Lushnikov, A. A. (1978). Coagulation in finite systems. *Journal of Colloid And Interface Science*, 65(2):276–285.

[Lushnikov, 2006] Lushnikov, A. A. (2006). Exact kinetics of sol-gel transition in a coagulating mixture. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 73(3):1–11.

[Lushnikov, 2014] Lushnikov, A. A. (2014). Composition distributions of particles in a gelling mixture. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 89(3):1–9.

[Man et al., 2014a] Man, S. M., Ekpenyong, A., Tourlomousis, P., Achouri, S., Cammarota, E., Hughes, K., Rizzo, A., Ng, G., Wright, J. A., Cicuta, P., Guck, J. R., and Bryant, C. E.

(2014a). Actin polymerization as a key innate immune effector mechanism to control Salmonella infection. *Proceedings of the National Academy of Sciences*, 111(49):17588–17593.

[Man et al., 2014b] Man, S. M., Hopkins, L. J., Nugent, E., Cox, S., Glück, I. M., Tourlomousis, P., Wright, J. A., Cicuta, P., Monie, T. P., and Bryant, C. E. (2014b). Inflammasome activation causes dual recruitment of NLRC4 and NLRP3 to the same macromolecular complex. *Proceedings of the National Academy of Sciences*, 111(20):7403–7408.

[Manji et al., 2002] Manji, G. A., Wang, L., Geddes, B. J., Brown, M., Merriam, S., Al-Garawi, A., Mak, S., Lora, J. M., Briskin, M., Jurman, M., Cao, J., Distefano, P. S., and Bertin, J. (2002). PYPAF1, a PYRIN-containing Apaf1-like protein that assembles with ASC and regulates activation of NF-κB. *Journal of Biological Chemistry*, 277(13):11570–11575.

[Marcus, 1968] Marcus, A. H. (1968). Stochastic Coalescence. *Technometrics*, 10(1):133–143.

[Marino and Kirschner, 2016] Marino, S. and Kirschner, D. E. (2016). A multi-compartment hybrid computational model predicts key roles for dendritic cells in tuberculosis infection. *Computation*, 4(4):1–25.

[Martin et al., 2014] Martin, B. N., Wang, C., Willette-Brown, J., Herjan, T., Gulen, M. F., Zhou, H., Bulek, K., Franchi, L., Sato, T., Alnemri, E. S., Narla, G., Zhong, X. P., Thomas, J., Klinman, D., Fitzgerald, K. A., Karin, M., Nuñez, G., Dubyak, G., Hu, Y., and Li, X. (2014). IKKα negatively regulates ASC-dependent inflammasome activation. *Nature Communications*, 5:4977.

[Martinon et al., 2002] Martinon, F., Burns, K., and Tschopp, J. (2002). The Inflammasome: A molecular platform triggering activation of inflammatory caspases and processing of proIL-β. *Molecular Cell*, 10(2):417–426.

[Masel et al., 1999] Masel, J., Jansen, V. A., and Nowak, M. A. (1999). Quantifying the kinetic parameters of prion replication. *Biophysical Chemistry*, 77:139–152.

[Mastroeni et al., 1991] Mastroeni, P., Arena, A., Costa, G. B., Liberto, M. C., Bonina, L., and Hormaeche, C. E. (1991). Serum TNFα in mouse typhoid and enhancement of a salmonella infection by anti-TNFα antibodies. *Microbial Pathogenesis*, 11(1):33–38.

[Mastroeni et al., 1996] Mastroeni, P., Harrison, J. A., Chabalgoity, J. A., and Hormaeche, C. E. (1996). Effect of interleukin 12 neutralization on host resistance and gamma interferon production in mouse typhoid. *Infection and Immunity*, 64(1):189–196.

[Mastroeni et al., 2000] Mastroeni, P., Vazquez-Torres, A., Fang, F. C., Xu, Y., Khan, S., Hormaeche, C. E., and Dougan, G. (2000). Antimicrobial actions of the nadph phagocyte oxidase and inducible nitric oxide synthase in experimental salmonellosis. II. Effects on microbial proliferation and host survival in vivo. *Journal of Experimental Medicine*, 192(2):237–247.

[Mastroeni et al., 1992] Mastroeni, P., Villarreal-Ramos, B., and Hormaeche, C. E. (1992). Role of T cells, TNFα and IFNγ in recall of immunity to oral challenge with virulent salmonellae in mice vaccinated with live attenuated aro- salmonella vaccines. *Microbial Pathogenesis*, 13(6):477–491.

[Mastroeni et al., 1993] Mastroeni, P., Villarreal-Ramos, B., and Hormaeche, C. E. (1993). Effect of late administration of anti-TNα antibodies on a Salmonella infection in the mouse model.

[Masumoto et al., 2001] Masumoto, J., Taniguchi, S., and Sagara, J. (2001). Pyrin N-terminal homology domain- and caspase recruitment domain-dependent oligomerization of ASC. *Biochemical and Biophysical Research Communications*, 280(3):652–655.

[Matsushita et al., 2009] Matsushita, K., Takeoka, M., Sagara, J., Itano, N., Kurose, Y., Nakamura, A., and Taniguchi, S. (2009). A splice variant of ASC regulates IL-1b release and aggregates differently from intact ASC. *Mediators of Inflammation*, 2009:1–6.

[Matthäus, 2006] Matthäus, F. (2006). Diffusion versus network models as descriptions for the spread of prion diseases in the brain. *Journal of Theoretical Biology*, 240(1):104–113.

[Matusiak et al., 2015] Matusiak, M., Van Opdenbosch, N., Vande Walle, L., Sirard, J.-C., Kanneganti, T.-D., and Lamkanfi, M. (2015). Flagellin-induced NLRC4 phosphorylation primes the inflammasome for activation by NAIP5. *Proceedings of the National Academy of Sciences of the United States of America*, 112(5):1541–1546.

[Matyszewski et al., 2021] Matyszewski, M., Zheng, W., Lueck, J., Mazanek, Z., Mohideen, N., Lau, A. Y., Egelman, E. H., and Sohn, J. (2021). Distinct axial and lateral interactions

within homologous filaments dictate the signaling specificity and order of the AIM2-ASC inflammasome. *Nature Communications*, 12(1):1–11.

[McDaniel et al., 2019] McDaniel, M., Keller, J. M., White, S., Baird, A., and Mcdaniel, M. (2019). A Whole-Body Mathematical Model of Sepsis Progression and Treatment Designed in the BioGears Physiology Engine. *Frontiers in Physiology*, 10:1321.

[Meng et al., 2009] Meng, G., Zhang, F., Fuss, I., Kitani, A., and Strober, W. (2009). A Mutation in the Nlrp3 Gene Causing Inflammasome Hyperactivation Potentiates Th17 Cell-Dominant Immune Responses. *Immunity*, 30(6):860–874.

[Méresse et al., 2001] Méresse, S., Unsworth, K. E., Habermann, A., Griffiths, G., Fang, F., Martínez-Lorenzo, M. J., Waterman, S. R., Gorvel, J. P., and Holden, D. W. (2001). Remodelling of the actin cytoskeleton is essential for replication of intravacuolar Salmonella. *Cellular Microbiology*, 3(8):567–577.

[Miao et al., 2003] Miao, E. A., Brittnacher, M., Haraga, A., Jeng, R. L., Welch, M. D., and Miller, S. I. (2003). Salmonella effectors translocated across the vacuolar membrane interact with the actin cytoskeleton. *Molecular Microbiology*, 48(2):401–415.

[Michaels et al., 2016] Michaels, T. C., Dear, A. J., Kirkegaard, J. B., Saar, K. L., Weitz, D. A., and Knowles, T. P. (2016). Fluctuations in the Kinetics of Linear Protein Self-Assembly. *Physical Review Letters*, 116(25):1–5.

[Michaels et al., 2017] Michaels, T. C., Liu, L. X., Meisl, G., and Knowles, T. P. (2017). Physical principles of filamentous protein self-assembly kinetics. *Journal of Physics Condensed Matter*, 29(15).

[Misawa et al., 2013] Misawa, T., Takahama, M., Kozaki, T., Lee, H., Zou, J., Saitoh, T., and Akira, S. (2013). Microtubule-driven spatial arrangement of mitochondria promotes activation of the NLRP3 inflammasome. *Nature Immunology*, 14(5):454–460.

[Mobilia et al., 2003] Mobilia, M., Krapivsky, P. L., and Redner, S. (2003). Kinetic anomalies in addition-aggregation processes. *Journal of Physics A: Mathematical and General*, 36(16):4533–4542.

[Moghaddas et al., 2018] Moghaddas, F., Zeng, P., Zhang, Y., Schützle, H., Brenner, S., Hofmann, S. R., Berner, R., Zhao, Y., Lu, B., Chen, X., Zhang, L., Cheng, S., Winkler,

S., Lehmberg, K., Canna, S. W., Czabotar, P. E., Wicks, I. P., De Nardo, D., Hedrich, C. M., Zeng, H., and Masters, S. L. (2018). Autoinflammatory mutation in NLRC4 reveals a leucine-rich repeat (LRR)–LRR oligomerization interface. *Journal of Allergy and Clinical Immunology*, 142(6):1956–1967.

[Moise and Friedman, 2021] Moise, N. and Friedman, A. (2021). A mathematical model of the multiple sclerosis plaque. *Journal of Theoretical Biology*, 512:110532.

[Morais et al., 2010] Morais, J. M., Papadimitrakopoulos, F., and Burgess, D. J. (2010). Biomaterials/tissue interactions: Possible solutions to overcome foreign body response. *AAPS Journal*, 12(2):188–196.

[Moriya et al., 2005] Moriya, M., Taniguchi, S., Wu, P., Liepinsh, E., Otting, G., and Sagara, J. (2005). Role of Charged and Hydrophobic Residues in the Oligomerization of the PYRIN Domain of ASC. *Biochemistry*, 44:575–583.

[Muñoz-Planillo et al., 2013] Muñoz-Planillo, R., Kuffa, P., Martínez-Colón, G., Smith, B. L., Rajendiran, T. M., and Núñez, G. (2013). K+ Efflux Is the Common Trigger of NLRP3 Inflammasome Activation by Bacterial Toxins and Particulate Matter. *Immunity*, 38(6):1142–1153.

[Muotiala and Mäkelä, 1990] Muotiala, A. and Mäkelä, P. H. (1990). The role of IFN-γ in murine Salmonella typhimurium infection. *Microbial Pathogenesis*, 8(2):135–141.

[Murakami et al., 2012] Murakami, T., Ockinger, J., Yu, J., Byles, V., McColl, A., Hofer, A. M., and Horng, T. (2012). Critical role for calcium mobilization in activation of the NLRP3 inflammasome. *Proceedings of the National Academy of Sciences of the United States of America*, 109(28):11282–11287.

[Murray, 1993] Murray, J. (1993). *Mathematical biology. II: Spatial models and biomedical applications*. Springer-Verlag, 3 edition.

[Murray and Hartley, 1935] Murray, R. C. and Hartley, G. S. (1935). Equilibrium between micelles and simple ions, with particular reference to the solubility of long-chain salts. *Transactions of the Faraday Society*, 31:183–189.

[Nagaraja et al., 2014] Nagaraja, S., Wallqvist, A., Reifman, J., and Mitrophanov, A. Y. (2014). Computational Approach To Characterize Causative Factors and Molecular Indicators of Chronic Wound Inflammation. *The Journal of Immunology*, 192:1824–1834.

[Nambayan et al., 2019] Nambayan, R. J. T., Sandin, S. I., Quint, D. A., Satyadi, D. M., and de Alba, E. (2019). The inflammasome adapter ASC assembles into filaments with integral participation of its two Death Domains, PYD and CARD. *Journal of Biological Chemistry*, 294(2):439–452.

[Nauciel and Espinasse-Maes, 1991] Nauciel, C. and Espinasse-Maes, F. (1991). Role of gamma interferon and tumor necrosis factor alpha in resistance to Salmonella typhimurium infection. *Infection and Immunity*, 60(2):450–454.

[Neuts, 1975] Neuts, M. F. (1975). Computational uses of the method of phases in the theory of queues. *Computers and Mathematics with Applications*, 1(2):151–166.

[Nnalue et al., 1992] Nnalue, N. A., Shnyra, A., Hultenby, K., and Lindberg, A. A. (1992). Salmonella choleraesuis and Salmonella typhimurium associated with liver cells after intravenous inoculation of rats are localized mainly in Kupffer cells and multiply intracellularly. *Infection and Immunity*, 60(7):2758–2768.

[Noureddini and Timm, 1992] Noureddini, H. and Timm, D. C. (1992). Kinetic Analysis of Competing Intramolecular and Intermolecular Polymerization Reactions. *Macromolecules*, 25(6):1725–1730.

[Nowak et al., 1998] Nowak, M. A., Krakauer, D. C., Klug, A., and May, R. M. (1998). Prion infection dynamics. *Integrative Biology*, 1(1):3–15.

[Ojcius et al., 2019] Ojcius, D. M., Jafari, A., Yeruva, L., Schindler, C. W., and Abdul-Sater, A. A. (2019). Dicer regulates activation of the NLRP3 inflammasome. *PLoS ONE*, 14(4):1–9.

[Olesen et al., 2005] Olesen, P., Ferkinghoff-Borg, J., Jensen, M. H., and Mathiesen, J. (2005). Diffusion, fragmentation, and coagulation processes: Analytical and numerical results. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 72(3):031103.

[Oosawa and Kasai, 1962] Oosawa, F. and Kasai, M. (1962). A theory of linear and helical aggregations of macromolecules. *Journal of Molecular Biology*, 4(1):10–21.

[Oroz et al., 2016] Oroz, J., Barrera-Vilarmau, S., Alfonso, C., Rivas, G., and de Alba, E. (2016). Asc Pyrin Domain Self-Associates and Binds Nlrp3 Using Equivalent Binding Interfaces. *Journal of Biological Chemistry*, 291(37):jbc.M116.741082.

[Paradis et al., 2021] Paradis, T., Bègue, H., Basmaciyan, L., Dalle, F., and Bon, F. (2021). Tight junctions as a key for pathogens invasion in intestinal epithelial cells. *International Journal of Molecular Sciences*, 22(5):1–21.

[Penner et al., 2012] Penner, K., Ermentrout, B., and Swigon, D. (2012). Pattern Formation in a Model of Acute Inflammation. *Journal of Applied Dynamical Systems*, 11(2):629–660.

[Pigozzo et al., 2012] Pigozzo, A. B., Macedo, G. C., Weber, R., and Lobosco, M. (2012). Computational Modeling of Microabscess Formation. *Computational and Mathematical Methods in Medicine*, page 736394.

[Platnich and Muruve, 2019] Platnich, J. M. and Muruve, D. A. (2019). NOD-like receptors and inflammasomes: A review of their canonical and non-canonical signaling pathways. *Archives of Biochemistry and Biophysics*, 670:4–14.

[Poh et al., 2008] Poh, J., Odendall, C., Spanos, A., Boyle, C., Liu, M., Freemont, P., and Holden, D. W. (2008). SteC is a Salmonella kinase required for SPI-2-dependent F-actin remodelling. *Cellular Microbiology*, 10(1):20–30.

[Pothen et al., 2015] Pothen, J. J., Poynter, M. E., and Bates, J. H. (2015). A computational model of unresolved allergic inflammation in chronic asthma. *American Journal of Physiology - Lung Cellular and Molecular Physiology*, 308(4):L384–L390.

[Price et al., 2015] Price, I., Mochan-keef, E. D., Swigon, D., Ermentrout, G. B., Lukens, S., Toapanta, F. R., Ross, T. M., and Clermont, G. (2015). The inflammatory response to influenza A virus (H1N1): An experimental and mathematical study. *Journal of Theoretical Biology*, 374:83–93.

[Proctor et al., 2013] Proctor, C. J., Boche, D., Gray, D. A., and Nicoll, J. A. (2013). Investigating Interventions in Alzheimer's Disease with Computer Simulation Models. *PLoS ONE*, 8(9).

[Proctor et al., 2012] Proctor, C. J., Pienaar, I. S., Elson, J. L., and Kirkwood, T. B. (2012). Aggregation, impaired degradation and immunization targeting of amyloid-beta dimers in

Alzheimers disease: A stochastic modelling approach. *Molecular Neurodegeneration*, 7(1).

[Proell et al., 2013] Proell, M., Gerlic, M., Mace, P., Reed, J., and Riedl, S. (2013). The CARD plays a critical role in ASC foci formation and inflammasome signaling. *Biochemical Journal*, 449(3):613–621.

[Pruppacher and Klett, 1978] Pruppacher, H. R. and Klett, J. D. (1978). *Microphysics of Clouds and Precipitation*. D. Reidel, Dortrecht.

[Prüss and Pujo-Menjouet, 2006] Prüss, J. and Pujo-Menjouet, L. (2006). Analysis of a model for the dynamics of prions. *Discrete and Continuous Dynamical Systems - Series B*, 6(1):225–235.

[Qu et al., 2012] Qu, Y., Misaghi, S., Izrael-Tomasevic, A., Newton, K., Gilmour, L. L., Lamkanfi, M., Louie, S., Kayagaki, N., Liu, J., Kömüves, L., Cupp, J. E., Arnott, D., Monack, D., and Dixit, V. M. (2012). Phosphorylation of NLRC4 is critical for inflammasome activation. *Nature*, 490(7421):539–542.

[R Core Team, 2016] R Core Team (2016). R: A language and environment for statistical computing.

[Ramabhadran et al., 1976] Ramabhadran, T. E., Peterson, T. W., and Seinfeld, J. H. (1976). Dynamics of aerosol coagulation and condensation. *AIChE Journal*, 22(5):840–851.

[Rescigno et al., 2001] Rescigno, M., Urbano, M., Valzasina, B., Francolini, M., Rotta, G., Bonasio, R., Granucci, F., Kraehenbuhl, J., and Ricciardi-Castagnoli, P. (2001). Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nature Immunology*, 2(4):361–7.

[Reynolds et al., 2006] Reynolds, A., Rubin, J., Clermont, G., Day, J., Vodovotz, Y., and Ermentrout, G. B. (2006). A reduced mathematical model of the acute inflammatory response : I. Derivation of model and analysis of anti-inflammation. *Journal of Theoretical Biology*, 242:220–236.

[Richter-Dahlfors et al., 1997] Richter-Dahlfors, A., Buchan, A. M., and Finlay, B. B. (1997). Murine salmonellosis studied by confocal microscopy: Salmonella typhimurium resides

intracellularly inside macrophages and exerts a cytotoxic effect on phagocytes in vivo. *Journal of Experimental Medicine*, 186(4):569–580.

[Robert and Sun, 2019] Robert, P. and Sun, W. E. N. (2019). On the asymptotic distribution of nucleation times of polymerization processes. *SIAM Journal on Applied Mathematics*, 79(5):2080–2106.

[Ross and Pompano, 2018] Ross, A. E. and Pompano, R. R. (2018). Diffusion of cytokines in live lymph node tissue using microfluidic integrated optical imaging. *Analytica Chimica Acta*, 1000:205–213.

[Rossi et al., 2019] Rossi, O., Coward, C., Goh, Y. S., Claassens, J. W., MacLennan, C. A., Verbeek, S. J., and Mastroeni, P. (2019). The essential role of complement in antibody-mediated resistance to Salmonella. *Immunology*, 156(1):69–73.

[Sahillioglu et al., 2014] Sahillioglu, A. C., Sumbul, F., Ozoren, N., and Haliloglu, T. (2014). Structural and Dynamics Aspects of ASC Speck Assembly. *Structure/Folding and Design*, 22(12):1722–1734.

[Salcedo et al., 2001] Salcedo, S. P., Noursadeghi, M., Cohen, J., and Holden, D. W. (2001). Intracellular replication of Salmonella typhimurium strains in specific subsets of splenic macrophages in vivo. *Cellular Microbiology*, 3(9):587–597.

[Salomonsky and Segal, 2017] Salomonsky, P.-m. and Segal, R. (2017). A mathematical system for human implantable wound model studies. *Letters in Biomathematics*, 4(1):77–100.

[Sanders et al., 2015] Sanders, M. G., Parsons, M. J., Howard, A. G. A., Liu, J., Fassio, S. R., Martinez, J. A., and Bouchier-Hayes, L. (2015). Single-cell imaging of inflammatory caspase dimerization reveals differential recruitment to inflammasomes. *Cell Death and Disease*, 6(7):e1813.

[Sanman et al., 2016] Sanman, L. E., Qian, Y., Eisele, N. A., Ng, T. M., van der Linden, W. A., Monack, D. M., Weerapana, E., and Bogyo, M. (2016). Disruption of glycolytic flux is a signal for inflammasome signaling and pyroptotic cell death. *eLife*, 5(MARCH2016):1–32.

[Sborgi et al., 2015] Sborgi, L., Ravotti, F., Dandey, V. P., Dick, M. S., Mazur, A., Reckel, S., Chami, M., Scherer, S., Huber, M., Böckmann, A., Egelman, E. H., Stahlberg, H., Broz, P., Meier, B. H., and Hiller, S. (2015). Structure and assembly of the mouse ASC inflammasome by combined NMR spectroscopy and cryo-electron microscopy. *Proceedings of the National Academy of Sciences*, 112(43):13237–13242.

[Sborgi et al., 2016] Sborgi, L., Rühl, S., Mulvihill, E., Pipercevic, J., Heilig, R., Stahlberg, H., Farady, C. J., Müller, D. J., Broz, P., and Hiller, S. (2016). GSDMD membrane pore formation constitutes the mechanism of pyroptotic cell death . *The EMBO Journal*, 35(16):1766–1778.

[Schaff et al., 1997] Schaff, J., Fink, C. C., Slepchenko, B., Carson, J. H., and Loew, L. M. (1997). A general computational framework for modeling cellular structure and function. *Biophysical Journal*, 73(3):1135–1146.

[Schmid-Burgk et al., 2016] Schmid-Burgk, J. L., Chauhan, D., Schmidt, T., Ebert, T. S., Reinhardt, J., Endl, E., and Hornung, V. (2016). A genome-wide CRISPR (clustered regularly interspaced short palindromic repeats) screen identifies NEK7 as an essential component of NLRP3 inflammasome activation. *Journal of Biological Chemistry*, 291(1):103–109.

[Schmidt et al., 2016] Schmidt, F. I., Lu, A., Chen, J. W., Ruan, J., Tang, C., Wu, H., and Ploegh, H. L. (2016). A single domain antibody fragment that recognizes the adaptor ASC defines the role of ASC domains in inflammasome assembly. *The Journal of Experimental Medicine*, 213(5):771–790.

[Schokker et al., 2013] Schokker, D., Bannink, A., Smits, M. A., and Rebel, J. M. (2013). A mathematical model representing cellular immune development and response to Salmonella of chicken intestinal tissue. *Journal of Theoretical Biology*, 330:75–87.

[Schreck and Yuan, 2013] Schreck, J. S. and Yuan, J. M. (2013). A kinetic study of amyloid formation: Fibril growth and length distributions. *Journal of Physical Chemistry B*, 117(21):6574–6583.

[Seal et al., 2011] Seal, J. B., Alverdy, J. C., Zaborina, O., and An, G. (2011). Agent-based dynamic knowledge representation of Pseudomonas aeruginosa virulence activation in

the stressed gut: Towards characterizing host-pathogen interactions in gut-derived sepsis. *Theoretical Biology and Medical Modelling*, 8(1):1–34.

[Seekhao et al., 2018] Seekhao, N., Shung, C., Jaja, J., and Mongeau, L. (2018). High-Performance Agent-Based Modeling Applied to Vocal Fold Inflammation and Repair. *Frontiers in Physiology*, 9(April):1–20.

[Sharif et al., 2019] Sharif, H., Wang, L., Wang, W. L., Magupalli, V. G., Andreeva, L., Qiao, Q., Hauenstein, A. V., Wu, Z., Núñez, G., Mao, Y., and Wu, H. (2019). Structural mechanism for NEK7-licensed activation of NLRP3 inflammasome. *Nature*, 570(7761):338–343.

[Sheikh et al., 2015] Sheikh, Z., Brooks, P. J., Barzilay, O., Fine, N., and Glogauer, M. (2015). Macrophages, foreign body giant cells and their response to implantable biomaterials. *Materials*, 8(9):5671–5701.

[Shen et al., 2019] Shen, C., Sharif, H., Xia, S., and Wu, H. (2019). Structural and mechanistic elucidation of inflammasome signaling by cryo-EM. *Current Opinion in Structural Biology*, 58:18–25.

[Sheppard et al., 2003] Sheppard, M., Webb, C., Heath, F., Mallows, V., Emilianus, R., Maskell, D., and Mastroeni, P. (2003). Dynamics of bacterial growth and distribution within the liver during Salmonella infection. *Cellular Microbiology*, 5(9):593–600.

[Shi et al., 2016a] Shi, H., Wang, Y., Li, X., Zhan, X., Tang, M., Fina, M., Su, L., Pratt, D., Hui Bu, C., Hildebrand, S., Lyon, S., Scott, L., Quan, J., Sun, Q., Russell, J., Arnett, S., Jurek, P., Chen, D., Kravchenko, V. V., Mathison, J. C., Moresco, E. M. Y., Monson, N. L., Ulevitch, R. J., and Beutler, B. (2016a). NLRP3 activation and mitosis are mutually exclusive events coordinated by NEK7, a new inflammasome component. *Nature Immunology*, 17(3):250–258.

[Shi et al., 2015a] Shi, J., Zhao, Y., Wang, K., Shi, X., Wang, Y., Huang, H., Zhuang, Y., Cai, T., Wang, F., and Shao, F. (2015a). Cleavage of GSDMD by inflammatory caspases determines pyroptotic cell death. *Nature*, 526(7575).

[Shi et al., 2016b] Shi, Z., Ben-Arieh, D., and Wu, C.-h. J. (2016b). A preliminary study of sepsis progression in an animal model using agent-based modeling. *International Journal of Modelling and Simulation*, 36(1-2):44–54.

[Shi et al., 2016c] Shi, Z., Chapes, S. K., Ben-Arieh, D., and Wu, C. H. (2016c). An agent-based model of a hepatic inflammatory response to salmonella: A computational study under a large set of experimental data. *PLoS ONE*, 11(8).

[Shi et al., 2015b] Shi, Z., Wu, C.-h. J., Ben-Arieh, D., and Simpson, S. Q. (2015b). Mathematical Model of Innate and Adaptive Immunity of Sepsis : A Modeling and Simulation Study of Infectious Disease. *BioMed Research International*, page 504259.

[Shimada et al., 2012] Shimada, K., Crother, T. R., Karlin, J., Dagvadorj, J., Chiba, N., Chen, S., Ramanujan, V. K., Wolf, A. J., Vergnes, L., Ojcius, D. M., Rentsendorj, A., Vargas, M., Guerrero, C., Wang, Y., Fitzgerald, K. A., Underhill, D. M., Town, T., and Arditi, M. (2012). Oxidized Mitochondrial DNA Activates the NLRP3 Inflammasome during Apoptosis. *Immunity*, 36(3):401–414.

[Shin et al., 2019] Shin, J. I., Lee, K. H., Joo, Y. H., Lee, J. M., Jeon, J., Jung, H. J., Shin, M., Cho, S., Kim, T. H., Park, S., Jeon, B. Y., Jeong, H., Lee, K., Kang, K., Oh, M., Lee, H., Lee, S., Kwon, Y., ho Oh, G., and Kronbichler, A. (2019). Inflammasomes and autoimmune and rheumatic diseases: A comprehensive review. *Journal of Autoimmunity*, 103(April):102299.

[Shirvani and van Roessel, 2004] Shirvani, M. and van Roessel, J. (2004). Spatially non-homogeneous coagulation equations with source terms. *Quarterly of Applied Mathematics*, 62(4):651–670.

[Siegmund-Schultze and Wagner, 2006] Siegmund-Schultze, R. and Wagner, W. (2006). Induced gelation in a two-site spatial coagulation model. *Annals of Applied Probability*, 16(1):370–402.

[Simonett and Walker, 2006] Simonett, G. and Walker, C. (2006). On the solvability of a mathematical model for prion proliferation. *Journal of Mathematical Analysis and Applications*, 324(1):580–603.

[Smoluchowski, 1916] Smoluchowski, M. (1916). Drei Vortrage uber Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen. *Zeitschrift für Physik*, 17:557–585.

[Solovyev et al., 2013] Solovyev, A., Mi, Q., Tzen, Y. T., Brienza, D., and Vodovotz, Y. (2013). Hybrid Equation/Agent-Based Model of Ischemia-Induced Hyperemia and Pressure Ulcer

Formation Predicts Greater Propensity to Ulcerate in Subjects with Spinal Cord Injury. *PLoS Computational Biology*, 9(5).

[Song et al., 2012] Song, L., Guo, Y., Deng, Q., and Li, J. (2012). TH17 functional study in severe asthma using agent based model. *Journal of Theoretical Biology*, 309:29–33.

[Srinivasula et al., 2002] Srinivasula, S. M., Poyet, J. L., Razmara, M., Datta, P., Zhang, Z., and Alnemri, E. S. (2002). The PYRIN-CARD protein ASC is an activating adaptor for caspase-1. *Journal of Biological Chemistry*, 277(24):21119–21122.

[Stanaway et al., 2019] Stanaway, J. D., Reiner, R. C., Blacker, B. F., Goldberg, E. M., Khalil, I. A., Troeger, C. E., Andrews, J. R., Bhutta, Z. A., Crump, J. A., Im, J., Marks, F., Mintz, E., Park, S. E., Zaidi, A. K. M., Abebe, Z., Abejie, A. N., Adedeji, I. A., Ali, B. A., Amare, A. T., Atalay, H. T., Avokpaho, E. F. G. A., Bacha, U., Barac, A., Bedi, N., Berhane, A., Browne, A. J., Chirinos, J. L., Chitheer, A., Dolecek, C., El Sayed Zaki, M., Eshrati, B., Foreman, K. J., Gemechu, A., Gupta, R., Hailu, G. B., Henok, A., Hibstu, D. T., Hoang, C. L., Ilesanmi, O. S., Iyer, V. J., Kahsay, A., Kasaeian, A., Kassa, T. D., Khan, E. A., Khang, Y.-H., Magdy Abd El Razek, H., Melku, M., Mengistu, D. T., Mohammad, K. A., Mohammed, S., Mokdad, A. H., Nachega, J. B., Naheed, A., Nguyen, C. T., Nguyen, H. L. T., Nguyen, L. H., Nguyen, N. B., Nguyen, T. H., Nirayo, Y. L., Pangestu, T., Patton, G. C., Qorbani, M., Rai, R. K., Rana, S. M., Ranabhat, C. L., Roba, K. T., Roberts, N. L. S., Rubino, S., Safiri, S., Sartorius, B., Sawhney, M., Shiferaw, M. S., Smith, D. L., Sykes, B. L., Tran, B. X., Tran, T. T., Ukwaja, K. N., Vu, G. T., Vu, L. G., Weldegebreal, F., Yenit, M. K., Murray, C. J. L., and Hay, S. I. (2019). The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Infectious Diseases*, 19(4):369–381.

[Stehlik et al., 2003a] Stehlik, C., Krajewska, M., Welsh, K., Krajewski, S., Godzik, A., and Reed, J. C. (2003a). The PAAD/PYRIN-only protein POP1/ASC2 is a modulator of ASC-mediated nuclear-factor-kappa B and pro-caspase-1 regulation. *The Biochemical journal*, 373:101–13.

[Stehlik et al., 2003b] Stehlik, C., Lee, S. H., Dorfleutner, A., Stassinopoulos, A., Sagara, J., and Reed, J. C. (2003b). Apoptosis-Associated Speck-Like Protein Containing a Caspase

Recruitment Domain Is a Regulator of Procaspase-1 Activation. *The Journal of Immunology*, 171(11):6154–6163.

[Stockmayer, 1943] Stockmayer, W. H. (1943). Theory of molecular size distribution and gel formation in branched-chain polymers. *The Journal of Chemical Physics*, 11(2):45–55.

[Swanson et al., 2019] Swanson, K. V., Deng, M., and Ting, J. P. (2019). The NLRP3 inflammasome: molecular activation and regulation to therapeutics. *Nature Reviews Immunology*, 19:477–489.

[Sweby, 1984] Sweby, P. K. (1984). High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws. *SIAM Journal on Numerical Analysis*, 21(5):995–1011.

[Tartey and Kanneganti, 2019] Tartey, S. and Kanneganti, T. D. (2019). Differential role of the NLRP3 inflammasome in infection and tumorigenesis. *Immunology*, 156(4):329–338.

[Tenthorey et al., 2020] Tenthorey, J. L., Chavez, R. A., Thompson, T. W., Deets, K. A., Vance, R. E., and Rauch, I. (2020). Nlrc4 inflammasome activation is nlrp3- And phosphorylation-independent during infection and does not protect from melanoma. *Journal of Experimental Medicine*, 217(7):1–9.

[Tenthorey et al., 2017] Tenthorey, J. L., Haloupek, N., López-Blanco, J. R., Grob, P., Adamson, E., Hartenian, E., Lind, N. A., Bourgeois, N. M., Chacón, P., Nogales, E., and Vance, R. E. (2017). The structural basis of flagellin detection by NAIP5: A strategy to limit pathogen immune evasion. *Science*, 358(6365):888–893.

[Tenthorey et al., 2014] Tenthorey, J. L., Kofoed, E. M., Daugherty, M. D., Malik, H. S., and Vance, R. E. (2014). Molecular Basis for Specific Recognition of Bacterial Ligands by NAIP/NLRC4 inflammasomes. *Molecular Cell*, 54(1):17–29.

[Thomas et al., 2009] Thomas, P. G., Dash, P., Aldridge, J. R., Ellebedy, A. H., Reynolds, C., Funk, A. J., Martin, W. J., Lamkanfi, M., Webby, R. J., Boyd, K. L., Doherty, P. C., and Kanneganti, T.-D. (2009). The Intracellular Sensor NLRP3 Mediates Key Innate and Healing Responses to Influenza A Virus via the Regulation of Caspase-1. *Immunity*, 30(4):566–575.

[Ting et al., 2008] Ting, J. P., Lovering, R. C., Alnemri, E. S., Bertin, J., Boss, J. M., Davis, B. K., Flavell, R. A., Girardin, S. E., Godzik, A., Harton, J. A., Hoffman, H. M., Hugot, J. P.,

Inohara, N., MacKenzie, A., Maltais, L. J., Nunez, G., Ogura, Y., Otten, L. A., Philpott, D., Reed, J. C., Reith, W., Schreiber, S., Steimle, V., and Ward, P. A. (2008). The NLR Gene Family: A Standard Nomenclature. *Immunity*, 28(3):285–287.

[Torres et al., 2019] Torres, M., Wang, J., Yannie, P. J., Shobha, G., Segal, R. A., and Reynolds, A. M. (2019). Identifying important parameters in the inflammatory process with a mathematical model of immune cell influx and macrophage polarization. *PLoS Computational Biology*, 15(7):e1007172.

[Trautmann, 1993] Trautmann, T. (1993). On the stochastic approach to the kinetics in polydisperse multicomponent populations. *Atmospheric Research*, 29(1-2):99–113.

[Turing, 1952] Turing, A. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 237:37–72.

[Vajjhala et al., 2014] Vajjhala, P. R., Kaiser, S., Smith, S. J., Ong, Q. R., Soh, S. L., Stacey, K. J., and Hill, J. M. (2014). Identification of multifaceted binding modes for pyrin and ASC pyrin domains gives insights into pyrin inflammasome assembly. *Journal of Biological Chemistry*, 289(34):23504–23519.

[Valeyev et al., 2010] Valeyev, N. V., Hundhausen, C., Umezawa, Y., and Kotov, N. V. (2010). A Systems Model for Immune Cell Interactions Unravels the Mechanism of Inflammation in Human Skin. *PLoS Computational Biology*, 6(12):e1001024.

[Van Dongen, 1987] Van Dongen, P. G. (1987). On the possible occurrence of instantaneous gelation in smoluchowski's coagulation equation. *Journal of Physics A: Mathematical and General*, 20(7):1889–1904.

[Vaughan et al., 2018] Vaughan, L. E., Ranganathan, P. R., Kumar, R. G., Wagner, A. K., and Rubin, J. E. (2018). A mathematical model of neuroinflammation in severe clinical traumatic brain injury. *Journal of Neuroinflammation*, 15(345).

[Vazquez-Torres et al., 1999] Vazquez-Torres, A., Jones-Carson, J., Bäumler, A. J., Falkow, S., Valdivia, R., Brown, W., Le, M., Berggren, R., Parks, W. T., and Fang, F. C. (1999). Extraintestinal dissemination CD18-expressing phagocytes. *Nature*, 401(October):623–626.

[Vazquez-Torres et al., 2000] Vazquez-Torres, A. A., Xu, Y., Jones-Carson, J., Holden, D. W., Lucia, S. M., Dinauer, M. C., Mastroeni, P., and Fang, F. C. (2000). Salmonella Salmonella Pathogenicity Pathogenicity Island Island 2-Dependent Evasion of the Phagocyte NADPH Oxidase. *Science*, 287(5458):1655–1658.

[Velloso et al., 2019] Velloso, F. J., Trombetta-Lima, M., Anschau, V., Sogayar, M. C., and Correa, R. G. (2019). NOD-like receptors: Major players (and targets) in the interface between innate immunity and cancer. *Bioscience Reports*, 29(4):1–21.

[Verhoef et al., 2005] Verhoef, P. A., Kertesy, S. B., Lundberg, K., Kahlenberg, J. M., and Dubyak, G. R. (2005). Inhibitory Effects of Chloride on the Activation of Caspase-1, IL-1β Secretion, and Cytolysis by the P2X7 Receptor. *The Journal of Immunology*, 175(11):7623–7634.

[Wagenmakers and Farrell, 2004] Wagenmakers, E. J. and Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin and Review*, 11(1):192–196.

[Warren et al., 2008] Warren, S. E., Mao, D. P., Rodriguez, A. E., Miao, E. A., and Aderem, A. (2008). Multiple Nod-Like Receptors Activate Caspase 1 during Listeria monocytogenes Infection. *The Journal of Immunology*, 180(11):7558 LP – 7564.

[Warsinske et al., 2016] Warsinske, H. C., Wheaton, A. K., Kim, K. K., Linderman, J. J., Moore, B. B., and Kirschner, D. E. (2016). Computational modeling predicts simultaneous targeting of fibroblasts and epithelial cells is necessary for treatment of pulmonary fibrosis. *Frontiers in Pharmacology*, 7(JUN).

[Wattis, 2006a] Wattis, J. A. (2006a). An introduction to mathematical models of coagulation-fragmentation processes: A discrete deterministic mean-field approach. *Physica D: Nonlinear Phenomena*, 222(1-2):1–20.

[Wattis, 2006b] Wattis, J. A. (2006b). Exact solutions for cluster-growth kinetics with evolving size and shape profiles. *Journal of Physics A: Mathematical and General*, 39(23):7283–7298.

[Wattis and King, 1998] Wattis, J. A. D. and King, J. R. (1998). Asymptotic solutions of the Becker-Döring equations Asymptotic solutions of the Becker – Döring equations. *Journal of Physics A: Mathematical and General*, 31:7169–7189.

[Williams and Iatropoulos, 2002] Williams, G. M. and Iatropoulos, M. J. (2002). Alteration of Liver Cell Function and Proliferation : Differentiation Between Adaption and Toxicity. *Toxic Pathology*, 30(1):41–53.

[Witten and Sander, 1981] Witten, T. A. and Sander, L. M. (1981). Diffusion-limited aggregation, a kinetic critical phenomenon. *Physical Review Letters*, 47(19):1400–1403.

[Woodward et al., 1995] Woodward, D. E., Tyson, R., Myerscough, M. R., Murray, J. D., Budrene, E. O., and Berg, H. C. (1995). Spatio-temporal patterns generated by Salmonella typhimurium. *Biophysical Journal*, 68(5):2181–2189.

[Yang et al., 2013] Yang, J., Zhao, Y., Shi, J., and Shao, F. (2013). Human NAIP and mouse NAIP1 recognize bacterial type III secretion needle protein for inflammasome activation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(35):14408–14413.

[Yang et al., 2018] Yang, X., Yang, F., Wang, W., Lin, G., Hu, Z., Han, Z., Qi, Y., Zhang, L., Wang, J., Sui, S. F., and Chai, J. (2018). Structural basis for specific flagellin recognition by the NLR protein NAIP5. *Cell Research*, 28:35–47.

[Yang et al., 2019] Yang, Y., Wang, H., Kouadir, M., Song, H., and Shi, F. (2019). Recent advances in the mechanisms of NLRP3 inflammasome activation and its inhibitors. *Cell Death and Disease*, 10(2).

[Yeh et al., 2017] Yeh, F. C., Liu, L., Hitchens, T. K., and Wu, Y. L. (2017). Mapping immune cell infiltration using restricted diffusion MRI. *Magnetic Resonance in Medicine*, 77(2):603–612.

[Yu et al., 2020] Yu, C., Du, F., Zhang, C., Li, Y., Liao, C., He, L., Cheng, X., and Zhang, X. (2020). Salmonella enterica serovar Typhimurium sseK3 induces apoptosis and enhances glycolysis in macrophages. *BMC Microbiology*, 20(1):1–9.

[Yvinec et al., 2016] Yvinec, R., Bernard, S., Hingant, E., and Pujo-Menjouet, L. (2016). First passage times in homogeneous nucleation: Dependence on the total number of particles. *Journal of Chemical Physics*, 144(3).

[Yvinec et al., 2012] Yvinec, R., Dorsogna, M. R., and Chou, T. (2012). First passage times in homogeneous nucleation and self-assembly. *Journal of Chemical Physics*, 137(24):1–16.

[Zamani et al., 2020] Zamani, P., Oskuee, R. K., Atkin, S. L., Navashenaq, J. G., and Sahebkar, A. (2020). MicroRNAs as important regulators of the NLRP3 inflammasome. *Progress in Biophysics and Molecular Biology*, 150:50–61.

[Zhang et al., 2015] Zhang, L., Chen, S., Ruan, J., Wu, J., Tong, A. B., Yin, Q., Li, Y., David, L., Lu, A., Wang, W. L., Marks, C., Ouyang, Q., Zhang, X., Mao, Y., and Wu, H. (2015). Cryo-EM structure of the activated NAIP2-NLRC4 inflammasome reveals nucleated polymerization. *Science*, 350(6259):404–409.

[Zhao et al., 2011] Zhao, Y., Yang, J., Shi, J., Gong, Y.-N., Lu, Q., Xu, H., Liu, L., and Shao, F. (2011). The NLRC4 inflammasome receptors for bacterial flagellin and type III secretion apparatus. *Nature*, 477(7366):596–600.

[Zhen and Zhang, 2019] Zhen, Y. and Zhang, H. (2019). NLRP3 inflammasome and inflammatory bowel disease. *Frontiers in Immunology*, 10(FEB):1–10.

[Zhou et al., 2011] Zhou, R., Yazdi, A. S., Menu, P., and Tschopp, J. (2011). A role for mitochondria in NLRP3 inflammasome activation. *Nature*, 469(7329):221–226.

[Ziraldo et al., 2015] Ziraldo, C., Solovyev, A., Allegretti, A., Krishnan, S., Henzel, M. K., Sowa, G. A., Brienza, D., An, G., Mi, Q., and Vodovotz, Y. (2015). A Computational, Tissue-Realistic Model of Pressure Ulcer Formation in Individuals with Spinal Cord Injury. *PLoS Computational Biology*, 11(6):1–28.